

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 09:42:01 ; Search time 83.31 Seconds
(without alignments)
8.000 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
X1 = Glu
X2 = Ala
X3 = Gly

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database :			
1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:	*	
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:	*	
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:	*	
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:	*	
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:	*	
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:	*	
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:	*	
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:	*	
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:	*	
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:	*	
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:	*	
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:	*	
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:	*	
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:	*	
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:	*	
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:	*	
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:	*	
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:	*	
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:	*	
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:	*	
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:	*	
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	19	90.5	47 19 AAW7585 Staphylococcus aur
2	19	90.5	81 22 AAU61772 Propionibacterium
3	19	90.5	97 22 AAU41551 Propionibacterium
4	19	90.5	100 22 ABC27830 Novel human diagno
5	19	90.5	113 22 AAU43242 Propionibacterium
6	19	90.5	117 22 AAU84257 Human immune/haema
7	19	90.5	118 20 AAY59973 Human endometrium
8	19	90.5	159 19 AAW31552 Collagen binding p
9	19	90.5	168 22 AAE11855 Staphylococcus aur
10	19	90.5	177 14 AAR39711 A. oryzae WT neutr
11	19	90.5	177 14 AAR39712 A. oryzae C6A neut

12	19	90.5	177	14	AAR39713	A. oryzae C78A neu
13	19	90.5	207	21	AAE24437	Human secreted pro
14	19	90.5	211	19	AAW31553	Collagen binding p
15	19	90.5	225	22	ABG03051	Novel human diagno
16	19	90.5	229	21	AG05085	Arabidopsis thalia
17	19	90.5	242	22	AAW25472	Human protein sequ
18	19	90.5	243	21	AG05084	Arabidopsis thalia
19	19	90.5	243	21	AAW5431	Human calcium chan
20	19	90.5	257	21	AG05083	Arabidopsis thalia
21	19	90.5	264	22	AAU68590	Human novel cytoki
22	19	90.5	270	22	ABE69721	Drosophila melanog
23	19	90.5	270	22	ABG03786	Novel human diagno
24	19	90.5	273	22	AAU31144	Novel human secret
25	19	90.5	290	22	AAU47789	Propionibacterium
26	19	90.5	305	22	ABG02081	Novel human diagno
27	19	90.5	323	22	AAW49464	Human protein sequ
28	19	90.5	330	11	AAW05528	High density lipop
29	19	90.5	332	21	AAW55898	Human myristoylate
30	19	90.5	332	21	AAW55899	Human myristoylate
31	19	90.5	333	21	AAW97409	Zebrafish Hsp-bind
32	19	90.5	341	22	AAW90789	C glutamicum prote
33	19	90.5	341	22	AAW79110	Corynebacterium gl
34	19	90.5	341	22	AAW79144	Corynebacterium gl
35	19	90.5	349	19	AAW44368	Aspergillus nidula
36	19	90.5	352	12	AAW14147	Pre-pro neutral pr
37	19	90.5	363	22	AAU36356	Pseudomonas aerugi
38	19	90.5	366	20	AAU05663	Maize caffeic O-me
39	19	90.5	377	21	AAW20561	Arabidopsis thalia
40	19	90.5	377	21	AAW41800	Arabidopsis thalia
41	19	90.5	386	21	AAW24042	Arabidopsis thalia
42	19	90.5	390	8	AAW70581	Protease biosynthe
43	19	90.5	393	20	AAW35147	Chlamydia pneumoni
44	19	90.5	422	20	AAW28643	Human serine prote
45	19	90.5	422	22	AAW74691	Human protease and
46	19	90.5	427	21	AAW41388	Arabidopsis thalia
47	19	90.5	431	21	AAW41387	Arabidopsis thalia
48	19	90.5	435	21	AAW24041	Arabidopsis thalia
49	19	90.5	440	21	AAW24040	Arabidopsis thalia
50	19	90.5	441	14	AAW31955	Sequence encoded b

ALIGNMENTS

RESULT 1
AAW77585
ID AAW77585 standard; Protein: 47 AA.

XX AC AAW77585;

XX AC AAW77585;

DT 30-OCT-1998 (first entry)

XX Staphylococcus aureus protein of unknown function.

XX Staphylococcus aureus protein; immune response induction; eye infection;
KW antibody production; T-cell immune response; gastrointestinal infection;
KW respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW central nervous system; kidney infection; urinary tract infection;
KW antimicrobial compound identification; broad spectrum antibiotic;
therapy.

XX Staphylococcus aureus.

XX Key Location/Qualifiers

FT Misc-difference 1..47 /note= "residues designated X are unspecified, and
represented as Xaa in the specification"

XX EP841394-A2.

XX 13-MAY-1998.

XX 24-SEP-1997; 97EP-0307485.

XX 24-SEP-1996; 96US-0027032.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 XX Black MT, Burnham MKR, Hodgson JE, Knowles DJC;
 PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;
 PI Ward JM;
 XX
 DR WPI: 1998-252940/23.
 DR N-PSDB; AAV53383.
 XX
 XX New nucleic acid sequences from Staphylococcus aureus WCHU29 -
 PT useful in vaccines and for treatment of bacterial infections of e.g.
 PT respiratory tract and central nervous system
 XX
 PS Claim 11; Page 267; 390pp; English.
 XX
 CC This sequence represents a Staphylococcus aureus protein of unknown
 CC function, and is encoded by a DNA sequence of the invention.
 CC The DNA sequences were isolated from Staphylococcus aureus WCHU29
 CC (NCIMB 40771). Host cells containing the DNA sequences are used to
 CC produce polypeptides or fragments. The proteins are used in the treatment
 CC of disease, for inducing an immune response by administering them, to
 CC produce antibody and/or T-cell immune response. Antagonists of the
 CC proteins are used for the inhibition of bacterial polypeptides.
 CC Conditions which may be treated include bacterial infections, especially
 CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,
 CC urinary tract, skin, bones and joints. The proteins can also be used to
 CC identify antimicrobial compounds which are broad spectrum antibiotics,
 CC especially useful in the treatment of H. pylori infection.
 XX
 SQ Sequence 47 AA;

Query Match 90.5%; Score 19; DB 19; Length 47;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 Db 31 eagats 36
 ||| |

RESULT 2
 AAU61772
 ID AAU61772 standard; Protein; 81 AA.
 AC AAU61772;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #23668.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59620.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 22967; 1069pp; English.
 XX
 CC Sequences AAU99105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 81 AA;

Query Match 90.5%; Score 19; DB 22; Length 81;
 Best Local Similarity 66.7%; Pred. No. 5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 Db 20 eagats 25
 ||| |

RESULT 3
 AAU41551
 ID AAU41551 standard; Protein; 97 AA.
 XX
 AC AAU41551;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #2447.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX

PA (CORI-) CORIXA CORP.
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59515.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 2746; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 97 AA;
 Query Match 90.5%; Score 19; DB 22; Length 97;
 Best Local Similarity 66.7%; Pred. No. 6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 eaqxqs 6
 ||| |
 Db 29 eagsas 34
 RESULT 4
 ABG27830
 ID ABG27830 standard; Protein; 100 AA.
 XX
 AC ABG27830;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #27821.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS92017.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 58189; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 100 AA;
 Query Match 90.5%; Score 19; DB 22; Length 100;
 Best Local Similarity 66.7%; Pred. No. 6.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 eaqxqs 6
 ||| |
 Db 95 eagats 100
 RESULT 5
 AAU43242
 ID AAU43242 standard; Protein; 113 AA.
 XX
 AC AAU43242;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #4138.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX

PA (CORI-) CORIXA CORP.
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59520.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX Example 1; SEQ ID No 4437; 1069pp; English.
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 113 AA;

Query Match 90.5%; Score 19; DB 22; Length 113;
 Best Local Similarity 66.7%; Pred. No. 6.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaggxs 6
 ||| |
 Db 99 eagtas 104

RESULT 6
 AAM84257
 ID AAM84257 standard; Protein: 117 AA.
 XX
 AC AAM84257;
 XX
 DT 07-NOV-2001 (first entry)
 DE
 DE Human immune/haematopoietic antigen SEQ ID NO:11850.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354..
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PA
 XX PI Rosen CA, Barash SC, Ruben SK;
 XX WPI; 2001-483426/52.
 XX N-PSDB; AAK57038.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX Claim 11; SEQ ID NO 11850; 3071pp + Sequence Listing; English.
 XX PS
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC

CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 117 AA;

Query Match 90.5%; Score 19; DB 22; Length 117;
 Best Local Similarity 66.7%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 ||| |
 Db 78 eaqtas 83

RESULT 7
 AAY59973
 ID AAY59973 standard; Protein; 118 AA.
 XX AC AAY59973;
 XX 31-JAN-2000 (first entry)
 DT Human endometrium tumour EST encoded protein 33.
 XX DE
 XX Endometrium; human; tumour; cancer; anticancer; cytostatic; EST:
 KW treatment; uterine; gene therapy; expressed sequence tag.
 XX OS Homo sapiens.
 XX PN DE19817948-A1.
 XX PD 21-OCT-1999.
 XX PF 17-APR-1998; 98DE-1017948.
 XX PR 17-APR-1998; 98DE-1017948.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX PA Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX WPI; 1999-591957/51.
 XX DR N-PSDB; AA241991.
 XX
 PT New nucleic acid sequences expressed in uterine cancer tissues, and
 PT derived polypeptides, for treatment of uterine and endometrial cancer
 PT and identification of therapeutic agents -
 XX
 XX Claim 23; Page 288; 444pp; German.
 CC This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B) i
 CC are used (i) to identify agents suitable for treatment of uterine or
 CC endometrial cancer; (ii) directly for treating these forms of cancer
 CC (including expression from gene therapy vectors) and (iii) for
 CC generation of specific antibodies. (A) are identified by assembling ESTs

CC (expressed sequence tags) from a particular tissue type before comparison
 CC of expression patterns. This allows a significantly longer fragment of
 CC the gene to be revealed, so should reduce the number of failures
 CC associated with the fact that ESTs from different libraries may represent
 CC different parts of the same unknown gene, distorting the estimated
 CC frequency of occurrence in a particular tissue. AAY5941-Y60328 represent
 CC protein fragments encoded by the human endometrium tumour cDNA library
 CC derived EST fragments represented in AAZ41981-242121.
 XX
 SQ Sequence 118 AA;

Query Match 90.5%; Score 19; DB 20; Length 118;
 Best Local Similarity 66.7%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 |||||
 Db 34 eagsas 39

RESULT 8
 AAW31552
 ID AAW31552 standard; Protein; 159 AA.
 AC AAW31552;

DT 21-MAY-1998 (first entry)

DE Collagen binding protein M17 epitope.

KW Collagen binding protein; cna gene; sepsis; infection;
 KW microbial surface component regionising adhesive matrix molecule;
 KW MSCRAMM; adhesin; vaccine; immunisation; diagnosis; therapy;
 KW epitope M17.

OS Staphylococcus aureus.

FH Key Location/Qualifiers
 FT Peptide 1..12
 FT /note= "vector pQE30-derived peptide"
 FT Protein 13..159
 FT /note= "epitope M17"

PN WO9743314-A2.

PD 20-NOV-1997.

PF 14-MAY-1997; 97WO-US08210.

PR 16-MAY-1996; 96US-0017678.

XX (UABR-) UAB RES FOUND.

PA (TEXA) UNIV TEXAS A & M SYSTEM.

PI Hook M, House-Pompeo K, Patti JM, Sthanam N, Symersky J;

DR WPI; 1998-008801/01.

DR N-PSDB; AAT93436.

XX Antibody that interacts with collagen binding domain of
 PT Staphylococcal cna gene product - useful to prevent bacterial sepsis
 PT in animal infected with Staphylococcus aureus

PS Claim 31; Page 114; 143pp; English.

XX This protein comprises Staphylococcus aureus collagen binding
 CC protein (CBP) epitope M17, i.e. amino acids 151-297 of full-length
 CC CBP, plus a vector-derived N-terminal peptide. Claimed 441, 849
 CC and 1500 bp nucleic acid sequences (see AAT93436-38) respectively
 CC encode CBP epitopes M17, M31 and M55 (see AAW31552-54) that confer
 CC protection against S. aureus infection. These nucleic acid
 CC sequences can be used in the recombinant production of the CBP

CC epitopes. The CBP protein and antigenic epitopes are contemplated
 CC for use in the treatment of pathological infections, especially to
 CC prevent bacterial adhesion to collagen. The claimed nucleic acids
 CC as well as claimed anti-CBP antibodies will also be of use in
 CC screening, diagnostic and therapeutic applications including active
 CC and passive immunisation and methods for the prevention of
 CC bacterial colonisation in an animal such as a human. The CBP
 CC epitopes are also contemplated for use in the preparation of
 CC vaccines and as carrier proteins in vaccine formulations, as well
 CC as in the formulation of compositions for the prevention of S.
 CC aureus infection.

SQ Sequence 159 AA;

Query Match 90.5%; Score 19; DB 19; Length 159;
 Best Local Similarity 66.7%; Pred. No. 9.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 |||||
 Db 28 eagtss 33

RESULT 9

AAE11855

ID AAE11855 standard; Protein; 168 AA.

AC AAE11855;

DT 18-DEC-2001 (first entry)

DE Staphylococcus aureus CNA19 protein.

KW Collagen-binding region; CNA19; Staphylococcus aureus infection;

KW Staphylococcus epidermidis infection; vaccine; CNA protein;

KW cross-reactive antibody.

OS Staphylococcus aureus.

FH Key Location/Qualifiers

FT Region 24..29
 FT /note= "Beta strand a; this region forms a part of
 FT the trench in the beta sheet"

FT Region 37..44

FT /note= "Beta strand b; this region forms a part of
 FT the trench in the beta sheet"

FT Region 55..61

FT /note= "Beta strand c"

FT Region 65..78

FT /note= "Beta strand d; a portion of this region forms
 FT a part of the trench in the beta sheet"

FT Region 82..84

FT /note= "Beta strand e; a portion of this region forms
 FT a part of the trench in the beta sheet"

FT Region 89..92

FT /note= "Alpha helix 1"

FT Region 93..96

FT /note= "Alpha helix 2"

FT Region 101..105

FT /note= "Beta strand f"

FT Region 110..115

FT /note= "Beta strand g"

FT Region 123..133

FT /note= "Beta strand h; a portion of this region forms
 FT a part of the trench in the beta sheet"

FT Region 140..149

FT /note= "Beta strand i"

FT Region 157..167

FT /note= "Beta strand j"

XX WO200170267-A1.

PN

PD 27-SEP-2001.
 PF 19-MAR-2001; 2001WO-US08554.
 XX
 PR 17-MAR-2000; 2000US-189868P.
 PR 25-APR-2000; 2000US-199370P.
 PR 15-AUG-2000; 2000US-225402P.
 XX
 XX (INHIT-) INHIBITEX INC.
 PA (TEXA) UNIV TEXAS A & M SYSTEM.
 PA (UYPA-) UNIV PAVIA.
 XX
 PI Hook M, Xu Y, Speziale P, Visai L, Casolini F, Patti J, Patel P;
 PI Domanski P;
 XX
 DR WPI; 2001-607512/69.
 XX
 PT Novel isolated antibody which recognizes collagen-binding peptide such
 PT as CNA19 peptide from *Staphylococcus aureus*, useful for preventing or
 PT treating *Staphylococcus aureus* or epidermidis infection -
 XX
 PS Example 2; Fig 2A; 107pp; English.
 XX
 CC The invention relates to an antibody which recognises a collagen-binding
 CC region including CNA19 of CNA protein from *Staphylococcus aureus*. This
 CC antibody is cross-reactive to collagen binding region of both *S. aureus*
 CC and *S. epidermidis*. It is useful for preventing or treating *S. aureus* or
 CC *S. epidermidis* infection in human or animal, and for displacing *S. aureus*
 CC or *S. epidermidis* bound to collagen. Antibody of the invention is useful
 CC for interfering with, modulating, and inhibiting the binding interactions
 CC between *Staphylococcus aureus* and collagen, for detecting the presence
 CC of *Staphylococcus aureus* or *Staphylococcus aureus* binding
 CC proteins, to diagnose *Staphylococcus aureus* infection, as research tools, for
 CC development of vaccine for passive immunisation against *Staphylococcus*
 CC infections, and in production facilities or laboratories to isolate
 CC additional quantities of collagen-binding proteins. It is also useful
 CC for treating medical instruments in order to reduce or eliminate the
 CC possibility of them becoming infected or further spreading the
 CC infection. The present sequence is *S. aureus* CNA19 protein.
 XX
 SQ Sequence 168 AA;

Query Match 90.5%; Score 19; DB 22; Length 168;
 Best Local Similarity 66.7%; Pred. No. 1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxss 6
 ||| |
 Db 16 eagtss 21

RESULT 10
 AAR39711
 ID AAR39711 standard; Protein; 177 AA.
 XX
 AC AAR39711;
 XX
 XX 24-JAN-1994 (first entry)
 XX
 XX A. oryzae WT neutral protease;
 XX
 KW Wildtype; protease; variant; neutral; mercapto group; food products;
 KW thermal stability; soy fermentation.
 XX
 OS Aspergillus oryzae.
 XX
 XX JP05168479-A.
 XX
 XX 02-JUL-1993.
 XX
 XX 26-DEC-1991; 91JP-0344443.
 XX
 XX 26-DEC-1991; 91JP-0344443.
 XX
 XX (SHSA) SHOKUHIN SANGYO KOSOKINO HENKA.
 XX
 DR WPI; 1993-247571/31.
 XX
 XX New variant neutral protease II - includes cysteine substd. with
 PT aminoacid having no mercapto gp. in aminoacid sequence of yellow
 PT green koji mould neutral protease II
 XX
 PS Disclosure; Page 8; 9pp; Japanese.
 XX
 CC The sequences given in AAR39712-13 represent variant neutral proteases
 CC based on the Aspergillus oryzae protease sequence (see also AAR39711).

PR 26-DEC-1991; 91JP-0344443.
 XX
 PA (SHSA) SHOKUHIN SANGYO KOSOKINO HENKA.
 XX
 DR WPI; 1993-247571/31.
 DR N-PSDB; AAQ46955.
 XX
 PT New variant neutral protease II - includes cysteine substd. with
 PT aminoacid having no mercapto gp. in aminoacid sequence of yellow
 PT green koji mould neutral protease II
 XX
 PS Disclosure; Page 7-8; 9pp; Japanese.
 XX
 CC This sequence represents a wildtype protease which may be used as the
 CC basis for the production of a variant neutral protease. The variant
 CC protease has either Cys6 or Cys78 substituted with an amino acid
 CC which has no -SH group. The variant proteases (see also AAR39713-14)
 CC have lower thermal stability than the WT and may be used in soy
 CC fermentation microorganisms. Soy produced by these microorganisms
 CC may be made into food products which will not be degraded by the
 CC presence of protease.
 XX
 SQ Sequence 177 AA;

Query Match 90.5%; Score 19; DB 14; Length 177;
 Best Local Similarity 66.7%; Pred. No. 1.1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxss 6
 ||| |
 Db 65 eagsts 70

RESULT 11
 AAR39712
 ID AAR39712 standard; Protein; 177 AA.
 XX
 AC AAR39712;
 XX
 XX 24-JAN-1994 (first entry)
 XX
 XX A. oryzae C6A neutral protease.
 XX
 KW Wildtype; protease; variant; neutral; mercapto group; food products;
 KW thermal stability; soy fermentation.
 XX
 OS Aspergillus oryzae.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 6
 FT /label= C6A
 XX
 XX JP05168479-A.
 XX
 XX 02-JUL-1993.
 XX
 XX 26-DEC-1991; 91JP-0344443.
 XX
 XX 26-DEC-1991; 91JP-0344443.
 XX
 XX (SHSA) SHOKUHIN SANGYO KOSOKINO HENKA.
 XX
 DR WPI; 1993-247571/31.
 XX
 XX New variant neutral protease II - includes cysteine substd. with
 PT aminoacid having no mercapto gp. in aminoacid sequence of yellow
 PT green koji mould neutral protease II
 XX
 PS Disclosure; Page 8; 9pp; Japanese.
 XX
 CC The sequences given in AAR39712-13 represent variant neutral proteases
 CC based on the Aspergillus oryzae protease sequence (see also AAR39711).

CC These variant proteases have either Cys6 or Cys78 substituted with an
 CC amino acid which has no -SH group. These variant proteases have lower
 CC thermal stability than the WT and may be used in soy fermentation
 CC microorganisms. Soy produced by these microorganisms may be made into
 CC food products which will not be degraded by the presence of protease.
 XX
 XX Sequence 177 AA;

Query Match 90.5%; Score 19; DB 14; Length 177;
 Best Local Similarity 66.7%; Pred. No. 1.1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 |||||
 Db 65 eagsts 70

RESULT 12
 AAR39713
 ID AAR39713 standard; Protein; 177 AA.

XX AC AAR39713;

XX DT 24-JAN-1994 (first entry)

XX DE A. oryzae C78A neutral protease.

XX KW Wildtype: protease; variant; neutral; mercapto group; food products;
 KW thermal stability; soy fermentation.

XX OS Aspergillus oryzae.

XX FH Key Location/Qualifiers
 FT Misc-difference 78
 FT /label= C78A

XX PN JP05168479-A.

XX PD 02-JUL-1993.

XX PF 26-DEC-1991; 91JP-0344443.

XX PR 26-DEC-1991; 91JP-0344443.

XX PA (SHSA) SHOKUHN SANGYO KOSOKINO HENKA.

XX PS WPI; 1993-247571/31.

XX PT New variant neutral protease II - includes cysteine substd. with
 PT aminoacid having no mercapto gp. in aminoacid sequence of yellow
 PT green koji mould neutral protease II

XX PS Disclosure; Page 8; 9pp; Japanese.

XX CC The sequences given in AAR39712-13 represent variant neutral proteases
 CC based on the Aspergillus oryzae protease sequence (see also AAR39711).
 CC These variant proteases have either Cys6 or Cys78 substituted with an
 CC amino acid which has no -SH group. These variant proteases have lower
 CC thermal stability than the WT and may be used in soy fermentation
 CC microorganisms. Soy produced by these microorganisms may be made into
 CC food products which will not be degraded by the presence of protease.
 XX

XX SQ Sequence 177 AA;

Query Match 90.5%; Score 19; DB 14; Length 177;
 Best Local Similarity 66.7%; Pred. No. 1.1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 |||||
 Db 65 eagsts 70

RESULT 13
 AAB24437

XX ID AAB24437 standard; Protein; 207 AA.

XX AC AAB24437;

XX DT 20-NOV-2000 (first entry)

XX DE Human secreted protein sequence encoded by gene 1 SEQ ID NO:62.

XX KW Human; secreted protein; cytostatic; antianaemic; antidiabetic;
 KW antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
 KW antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
 KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
 KW immune system disorder; angiogenesis; hyperproliferative disorder;
 KW cardiovascular disorder; apoptosis; hyperproliferative disorder;
 KW infectious disease; wound healing.

XX OS Homo sapiens.

XX PN WO200035937-A1.

XX PD 22-JUN-2000.

XX PF 16-DEC-1999; 99WO-US29950.

XX PR 17-DEC-1998; 98US-0112809.

XX PR 18-DEC-1998; 98US-0113006.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
 PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;

XX DR WPI; 2000-431566/37.

XX DR N-PSDB; AAA78381.

XX PT Forty seven human nucleic acids encoding secreted proteins, useful in
 PT the treatment, prevention and diagnosis of cancers, disorders of the
 PT immune system, angiogenesis disorders, neurological diseases and
 PT hyperproliferative disorders -

XX PS Claim 11; Page 478-479; 562pp; English.

XX CC The polynucleotide sequence given in AAA78381 to AAA78432 encode the
 CC human secreted proteins given in AAB24437 to AAB24604. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic; antianaemic;
 CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
 CC antiarthritic; antipsoriatic; antiangiogenic; cardiant; anti-HIV;
 CC nootropic; neuroprotective; antimicrobial and antiparkinsonian.
 CC Human secreted protein polynucleotides, polypeptides, antagonists and/or
 CC agonists may be useful in treating, preventing, and/or diagnosing other
 CC diseases, disorders, and/or conditions such as: (a) cancers; (b)
 CC disorders of the immune system; (c) angiogenesis disorders; (d)
 CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
 CC associated with increase apoptosis; (g) neurological diseases; and
 CC (h) infectious diseases. They are also used to promote wound healing.
 CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the
 CC exemplification of the present invention.

XX SQ Sequence 207 AA;

Query Match 90.5%; Score 19; DB 21; Length 207;
 Best Local Similarity 66.7%; Pred. No. 1.3e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 |||||
 Db 77 eagats 82

RESULT 14
AAW31553
ID AAW31553 standard; Protein; 211 AA.
XX AC AAW31553;
XX DT 21-MAY-1998 (first entry)
XX DE Collagen binding protein M31 epitope.
XX KW Collagen binding protein; cna gene; sepsis; infection;
KW microbial surface component recognising adhesive matrix molecule;
KW MSCRAMM; adhesin; vaccine; immunisation; diagnosis; therapy;
KW epitope M31.
XX OS Staphylococcus aureus.
XX FH Key Location/Qualifiers
FT Peptide 1..12
FT Protein /note= "vector pQE30-derived peptide"
FT Protein 13..211
FT Protein /note= "epitope M31"
XX WO9743314-A2.
XX PN 20-NOV-1997.
XX PD 14-MAY-1997; 97WO-US08210.
XX PF 16-MAY-1996; 96US-0017678.
XX PR (UABR-) UAB RES FOUND.
XX PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX PI Hook M, House-Pompeo K, Patti JM, Sthanam N, Symersky J;
XX DR WPI; 1998-008801/01.
XX DR N-PSDB; AAT93437.
XX PT Antibody that interacts with collagen binding domain of
PT Staphylococcal cna gene product - useful to prevent bacterial sepsis
PT in animal infected with Staphylococcus aureus
XX PS Claim 31; Page 115-116; 143pp; English.
XX CC This protein comprises Staphylococcus aureus collagen binding
CC protein (CBP) epitope M31, i.e. amino acids 61-343 of full-length
CC CBP, plus a vector-derived N-terminal peptide. Claimed 441, 849
CC and 1500 bp nucleic acid sequences (see AAT93436-38) respectively
CC encode CBP epitopes M17, M31 and M55 (see AAW31552-54) that confer
CC protection against S. aureus infection. These nucleic acid
CC sequences can be used in the recombinant production of the CBP
CC epitopes. The CBP protein and antigenic epitopes are contemplated
CC for use in the treatment of pathological infections, especially to
CC prevent bacterial adhesion to collagen. The claimed nucleic acids
CC as well as claimed anti-CBP antibodies will also be of use in
CC screening, diagnostic and therapeutic applications including active
CC and passive immunisation and methods for the prevention of
CC bacterial colonisation in an animal such as a human. The CBP
CC epitopes are also contemplated for use in the preparation of
CC vaccines and as carrier proteins in vaccine formulations, as well
CC as in the formulation of compositions for the prevention of S.
CC aureus infection.
XX SQ Sequence 211 AA;

Query Match 90.5%; Score 19; DB 19; Length 211;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxss 6
Db 34 eagtss 39
RESULT 15
ABG03051
ID ABG03051 standard; Protein; 225 AA.
XX AC ABG03051;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #3042.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS67238.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX PS Claim 20; SEQ ID No 33410; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 225 AA;

Query Match 90.5%; Score 19; DB 22; Length 225;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxss 6

Db ||| |
16 eagtss 21

RESULT 16
AAG05085
ID AAG05085 standard; Protein; 229 AA.
XX AC AAG05085;
DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 1356.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

Best Local Similarity 66.7%; Pred. No. 1.5e+03;				
Matches	4; Conservative	0; Mismatches	2; Indels	0; Gaps
QY 1 eaqxxx 6				
Db	203 eagsss	208		
RESULT 18				
AAG05084				
ID	AAG05084 standard; Protein; 243 AA.			
XX				
AC	AAG05084;			
XX				
DT	17-OCT-2000 (first entry)			
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 1355.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
KW	termination sequence.			
XX				
OS	Arabidopsis thaliana.			
XX				
PN	EP1033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
PR	25-FEB-1999;	99US-0121825.	16-JUN-1999;	99US-0139452.
PR	05-MAR-1999;	99US-0123180.	16-JUN-1999;	99US-0139453.
PR	09-MAR-1999;	99US-0123548.	17-JUN-1999;	99US-0139492.
PR	23-MAR-1999;	99US-0125788.	18-JUN-1999;	99US-0139454.
PR	25-MAR-1999;	99US-0126264.	18-JUN-1999;	99US-0139455.
PR	29-MAR-1999;	99US-0126785.	18-JUN-1999;	99US-0139456.
PR	01-APR-1999;	99US-0127462.	18-JUN-1999;	99US-0139457.
PR	06-APR-1999;	99US-0128234.	18-JUN-1999;	99US-0139458.
PR	08-APR-1999;	99US-0128714.	18-JUN-1999;	99US-0139459.
PR	16-APR-1999;	99US-0129845.	18-JUN-1999;	99US-0139460.
PR	19-APR-1999;	99US-0130077.	18-JUN-1999;	99US-0139461.
PR	21-APR-1999;	99US-0130449.	18-JUN-1999;	99US-0139462.
PR	23-APR-1999;	99US-0130510.	18-JUN-1999;	99US-0139463.
PR	28-APR-1999;	99US-0130891.	18-JUN-1999;	99US-0139750.
PR	30-APR-1999;	99US-0131449.	18-JUN-1999;	99US-0139763.
PR	04-MAY-1999;	99US-0132048.	21-JUN-1999;	99US-0139817.
PR	05-MAY-1999;	99US-0132484.	22-JUN-1999;	99US-0139899.
PR	06-MAY-1999;	99US-0132485.	23-JUN-1999;	99US-0140353.
PR	06-MAY-1999;	99US-0132486.	23-JUN-1999;	99US-0140354.
PR	07-MAY-1999;	99US-0132487.	24-JUN-1999;	99US-0140695.
PR	11-MAY-1999;	99US-0132863.	28-JUN-1999;	99US-0140823.
PR	14-MAY-1999;	99US-0134256.	29-JUN-1999;	99US-0140991.
PR	14-MAY-1999;	99US-0134218.	30-JUN-1999;	99US-0141287.
PR	14-MAY-1999;	99US-0134219.	01-JUL-1999;	99US-0141842.
PR	14-MAY-1999;	99US-0134221.	01-JUL-1999;	99US-0142154.
PR	14-MAY-1999;	99US-0134370.	02-JUL-1999;	99US-0142055.
PR	18-MAY-1999;	99US-0134768.	06-JUL-1999;	99US-0142390.
PR	19-MAY-1999;	99US-0134941.	08-JUL-1999;	99US-0142803.
PR	20-MAY-1999;	99US-0135124.	09-JUL-1999;	99US-0142920.
PR	21-MAY-1999;	99US-0135353.	12-JUL-1999;	99US-0142977.
PR	25-MAY-1999;	99US-0135629.	13-JUL-1999;	99US-0143542.
PR	25-MAY-1999;	99US-0136021.	14-JUL-1999;	99US-0143624.
PR	27-MAY-1999;	99US-0136392.	15-JUL-1999;	99US-0144005.
PR	28-MAY-1999;	99US-0136782.	16-JUL-1999;	99US-0144085.
PR	01-JUN-1999;	99US-0137222.	16-JUL-1999;	99US-0144086.
PR	03-JUN-1999;	99US-0137528.	19-JUL-1999;	99US-0144325.
PR	04-JUN-1999;	99US-0137502.	19-JUL-1999;	99US-0144331.
PR	07-JUN-1999;	99US-0137724.	19-JUL-1999;	99US-0144332.
PR	08-JUN-1999;	99US-0138094.	19-JUL-1999;	99US-0144334.
PR	10-JUN-1999;	99US-0138540.	19-JUL-1999;	99US-0144335.
PR	10-JUN-1999;	99US-0138847.	20-JUL-1999;	99US-0144352.
PR	14-JUN-1999;	99US-0139119.	20-JUL-1999;	99US-0144632.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 243;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxgs 6

Db 170 eagsss 175

RESULT 19
AAY95431
ID AAY95431 standard; Protein; 243 AA.

XX
AC AAY95431;

XX
DT 10-OCT-2000 (first entry)

XX
DE Human calcium channel SOC-2/CRAC-1.

XX
KW SOC-2/CRAC-1; calcium channel; human; store operated channel;
calcium release activated channel; therapy; diagnosis;
lymphocyte proliferative disorder.

XX
OS Homo sapiens.

XX
Key Location/Qualifiers

FT Misc-difference 103 /note= "encoded by SCA"

FT Misc-difference 104 /note= "encoded by CSA"

FT Misc-difference 105 /note= "encoded by RSC"

FT Misc-difference 109 /note= "encoded by GNT"

FT Misc-difference 141 /note= "encoded by NCA"

XX
WO2000040614-A2.

XX
13-JUL-2000.

XX
20-DEC-1999; 99WO-US29996.

XX
30-DEC-1998; 98US-0114220.

XX
29-JAN-1999; 99US-0120018.

XX
22-JUN-1999; 99US-0140415.

XX
(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX
Scharenberg AM;

XX
WPI; 2000-465957/40.

XX
N-PSDB; AAA49918.

XX
New SOC/CRAC calcium channel polynucleotides and polypeptides used to
diagnose and treat proliferative disorders associated with the channel,
and to screen for novel modulators of the channel -

XX
Claim 14; Page 58-59; 108pp; English.

XX
The present sequence is that of a partial sequence of human
SOC-2/CRAC-1 (full-length sequence given in AAY95435), as deduced
from a partial cDNA clone (see AAA49918). SOC-2/CRAC-1 is a member
of a novel family of store operated channel (SOC) or calcium release
activated channel (CRAC) polypeptides that modulate Ca²⁺ flux into
and out of a cell, and which may be activated upon depletion of
Ca²⁺ from intracellular calcium stores, allowing Ca²⁺ influx into
a cell. SOC-2/CRAC-1 is expressed predominantly in human
haematopoietic cells, liver, spleen, heart and kidney.
Compositions for expressing functional SOC/CRAC calcium channel
polypeptides in cells are expected to be useful for treating
patients that have reduced extracellular calcium influx into their
SOC/CRAC-expressing cells. They will also be useful for delivering
therapeutic and/or imaging agents to such cells to modulate
proliferation and growth. SOC/CRAC polypeptides also represent
targets for designing and/or identifying inhibitors that block
lymphocyte proliferation and binding agents that selectively bind
to SOC/CRAC polypeptides to which drugs or toxins can be conjugated
for delivery to SOC/CRAC expressing cells. Methods for determining
the level of SOC/CRAC expression in a subject can be used to assess
the presence, or absence, or stage of a proliferative disorder,

CC e.g. a lymphocyte proliferative disorder.
XX Sequence 243 AA;
SQ

Query Match 90.5%; Score 19; DB 21; Length 243;
Best Local Similarity 66.7%; Pred. NO. 1.5e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 53 eagss 58

RESULT 20
AAG05083
ID AAG05083 standard; Protein; 257 AA.
XX AC AAG05083;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 1354.
XX DE
XX DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD
XX PF 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PF 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 257;
 Best Local Similarity 66.7%; P.fed. No. 1.6e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 eagxxs 6
 ||| |
 Db 184 eagsss 189
 RESULT 21
 AAU68590
 ID AAU68590 standard; Protein; 264 AA.
 XX
 AC AAU68590;
 XX
 DT 16-JAN-2002 (first entry)
 XX
 DE Human novel cytokine encoded by cDNA 790CIP2D_8 #2.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation;
 KW antinflammatory; stem cell growth factor; activin; inhibin; cancer;
 KW nervous system disease; neuropathy; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; spinal cord disorder;
 KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopaenia; stem cell disorder;
 KW aplastic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoporosis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200175093-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US10484.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR 22-SEP-2000; 2000US-0668680.
 PR 23-OCT-2000; 2000US-0695618.
 PR 30-NOV-2000; 2000US-0728711.
 PR 14-MAR-2001; 2000US-0728711.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;
 PI Yang Y, Zhao QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-626432/72.
 DR N-PSDB; AAS59882.
 XX
 PT New polypeptides and nucleic acids, useful for diagnosis, treatment of
 PT inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
 PT degenerative disorders, cancer and promoting wound healing
 XX
 PS Claim 20; Page 328; 336pp; English.
 XX
 CC The invention relates to isolated human polypeptides (which may be
 CC cytokines) and the polynucleotides encoding them. The protein is useful
 CC for identifying a compound which binds to it (e.g. modulators, agonists
 CC and antagonists). The polynucleotides are useful as an array for mismatch
 CC detection. The proteins and nucleic acids are useful as nutritional
 CC sources or supplements. The protein exhibits activity relating
 CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity, immune stimulating or immune
 CC suppressing and activin or inhibin related activities. The proteins (and
 CC antibodies raised against them) and nucleic acids are therefore useful in
 CC the diagnosis and treatment of diseases and disorders such as cancer,
 CC central and peripheral nervous system diseases and neuropathies,
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, and in tissue repair, healing of burns, incisions, ulcers, for
CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or
CC periodontal disease, lung or liver fibrosis, reperfusion injury in
CC various tissues, various immune deficiencies and disorders including
CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
CC such as asthma or other respiratory problems, coagulation disorders,
CC haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory
CC bowel disease, viral infection and are useful in altering bodily
CC characteristics. The present sequence represents a novel protein of the
CC invention.

XX Sequence 264 AA;

Query Match 90.5%; Score 19; DB 22; Length 264;
Best Local Similarity 66.7%; Pred No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 207 eagaas 212

RESULT 22

ABB69721
ID ABB69721 standard; Protein; 270 AA.

XX ABB69721;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 35955.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

DR N-PSDB; ABL13824.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PS Disclosure; SEQ ID NO 35955; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 270 AA;

Query Match 90.5%; Score 19; DB 22; Length 270;
Best Local Similarity 66.7%; Pred No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 169 eagsts 174

RESULT 23

ABG03786

ID ABG03786 standard; Protein; 270 AA.

XX ABG03786;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #3777.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS67973.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 34145; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published/pct_sequences.

XX SQ Sequence 270 AA;

Query Match 90.5%; Score 19; DB 22; Length 270;
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 ||| |
 Db 16 eagtss 21

RESULT 24

AAU31144
 ID AAU31144 standard; Protein; 273 AA.

XX AC AAU31144;

XX DT 18-DRC-2001 (first entry)

XX DE Novel human secreted protein #1635.

KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US08656.

XX PR 18-APR-2000; 2000US-0552929.

XX PR 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -

PS Claim 20; Page 409; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX SQ Sequence 273 AA;

Query Match 90.5%; Score 19; DB 22; Length 273;
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 ||| |
 Db 35 eagass 40

RESULT 25

AAU47789

ID AAU47789 standard; Protein; 290 AA.

XX AC AAU47789;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #8685.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59539.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

PS Example 1; SEQ ID No 8984; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published/pct_sequences.

XX SQ Sequence 290 AA;

Query Match 90.5%; Score 19; DB 22; Length 290;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 79 eagass 84

RESULT 26

ABG02081
ID ABG02081 standard; Protein; 305 AA.

XX AC ABG02081;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #2072.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS66268.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID No 32440; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX CC sequence 305 AA;

XX SQ

Query Match 90.5%; Score 19; DB 22; Length 305;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 178 eagass 183

RESULT 27

AB94964
ID AAB94964 standard; Protein; 323 AA.

XX AC AAB94964;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:16523.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX KW Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
XX PT full-length cDNAs defined in the specification, and for the detection
XX PT and/or diagnosis of the abnormality of the proteins encoded by the
XX PT full-length cDNAs -

XX PS Claim 8; SEQ ID 16523; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602
XX CC full-length cDNAs defined in the specification. Where a primer set
XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX CC to the complementary strand of a polynucleotide which comprises one of
XX CC the 5602 nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in
XX CC the specification. The primer sets can be used in antisense therapy and
XX CC in gene therapy. The primers are useful for synthesising polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX CC represent oligonucleotides, all of which are used in the exemplification
XX CC of the present invention.

XX CC sequence 323 AA;

XX SQ

Query Match 90.5%; Score 19; DB 22; Length 323;
 Best Local Similarity 66.7%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 308 eagsts 313

RESULT 28
 AAR05528
 ID AAR05528 standard; protein; 330 AA.
 XX
 AC AAR05528;
 XX
 DT 23-OCT-1990 (first entry)
 XX
 DE High density lipoprotein (HDL) binding protein.
 XX
 KW High density lipoprotein; HDL-binding protein; atherosclerosis;
 KW hypercholesterolaemia; ds.
 XX
 OS Homo sapiens.
 XX
 PN W09005744-A.
 XX
 PD 31-MAY-1990.
 XX
 PF 17-NOV-1989; 89WO-0005169.
 XX
 PR 18-NOV-1988; 88US-0273388.
 XX
 PA (UNIV) UNIV OF WASHINGTON.
 XX
 XX (ZYMO-) ZYMOGENETICS INC.
 XX
 PI Oram JF, McKnight GL, Hart CE, Curtis DA;
 XX
 XX WPI; 1990-193405/25.
 DR N-PSDB; AAQ04784.
 XX
 PT New mammalian proteins binding high density lipoprotein sub-class 3 -
 PT DNA encoding them and derived antibodies, for screening
 PT potentially therapeutic HDL analogues and for diagnosing risk of
 PT atherosclerosis.
 XX
 PS Claim 4; Fig 1A-D; 79pp; English.
 XX
 CC The protein product may be used to raise Abs, and the cDNA to
 CC create probes, both useful in screening for HDL analogues,
 CC agonists and antagonists, and in identifying abnormalities in the
 CC HDL binding/receptor pathway. HDL analogues can be used in treating
 CC hypercholesterolaemia and atherosclerosis
 XX
 SQ Sequence 330 AA;

Query Match 90.5%; Score 19; DB 11; Length 330;
 Best Local Similarity 66.7%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 205 eagas 210

RESULT 29
 AAY95898
 ID AAY95898 standard; Protein; 332 AA.
 XX
 AC AAY95898;
 XX

DT 20-NOV-2000 (first entry)
 XX Human myristoylated alanine-rich C kinase substrate MARCKS.
 DE
 XX MARCKS; myristoylated alanine-rich C kinase substrate; human;
 KW mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
 KW chronic obstructive pulmonary disease; pneumonia; emphysema;
 KW influenza; rhinitis; therapy.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 84 /note- "Ser in sequence of AAY95899"
 FT Misc-difference 119 /note- "Ala in sequence of AAY95899"
 FT Peptide 2..25
 FT /note- "MANS peptide of AAY95898"
 FT Peptide 152..176
 FT /note- "MA-PSD peptide of AAY95897"
 XX
 PN W0200050062-A2.
 XX
 XX 31-AUG-2000.
 PD
 XX 24-FEB-2000; 2000WO-US05050.
 PF
 XX 24-FEB-1999; 99US-0256154.
 PR
 XX (UYNC-) UNIV NORTH CAROLINA STATE.
 PA
 XX Li Y, Martin LD, Adler KB;
 XX
 XX WPI; 2000-572036/53.
 DR N-PSDB; AAA50339.
 XX
 PT Regulating mucus secretion by a mucus-secreting cell, useful for
 PT treating e.g. bronchitis, asthma or pneumonia, by administering a
 PT compound that inhibits or enhances myristoylated alanine-rich C-kinase
 PT substrate protein -
 XX
 PS Claim 3; Page 42-43; 66pp; English.
 XX
 CC The present sequence is that of human myristoylated alanine-rich C
 CC kinase substrate MARCKS protein, a major cellular substrate. The
 CC invention relates to methods of inhibiting mucus secretion by a
 CC mucus-secreting cell by administering a compound that inhibits
 CC MARCKS protein-related mucus secretion. Such compounds include
 CC active fragments of MARCKS protein such as MANS peptide (see
 CC AAY95897) and MA-PDS peptide (see AAY95897), which corresponds to a
 CC phosphorylation site of MARCKS. The inhibitor compounds can be
 CC used to treat conditions such as bronchitis, cystic fibrosis,
 CC chronic obstructive pulmonary disease, asthma, emphysema,
 CC pneumonia, influenza, rhinitis and the common cold. An alternative
 CC sequence for MARCKS is provided in AAY95899, which differs from the
 CC present sequence at 2 amino acid residues, Ala-84 (Ser) and
 CC Pro-119 (Ala).
 XX
 SQ Sequence 332 AA;

Query Match 90.5%; Score 19; DB 21; Length 332;
 Best Local Similarity 66.7%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 206 eagas 211

RESULT 30
 AAY95899
 ID AAY95899 standard; Protein; 332 AA.

XX AC AAY95899;
 XX XX 20-NOV-2000 (first entry)
 XX XX Human myristoylated alanine-rich C kinase substrate MARCKS.
 XX DE
 XX KW MARCKS; myristoylated alanine-rich C kinase substrate; human;
 KW mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
 KW chronic obstructive pulmonary disease; pneumonia; emphysema;
 KW influenza; rhinitis; therapy.
 XX KW Homo sapiens.
 OS XX
 FH Key Location/Qualifiers
 FT Misc-difference 84 /note= "Ala in sequence of AAY95898"
 FT FT
 FT Misc-difference 119 /note= "Pro in sequence of AAY95898"
 FT FT
 FT Peptide 2..25
 FT /note= "MANS peptide of AAY95896"
 FT 152..176
 FT /note= "NA-PSD peptide of AAY95897"
 FT FT
 XX WO200050062-A2.
 XX 31-AUG-2000.
 XX 24-FEB-2000; 2000WO-US05050.
 XX 24-FEB-1999; 99US-0256154.
 XX (UYN-) UNIV NORTH CAROLINA STATE.
 XX LI Y, Martin LD, Adler KB;
 XX WPI: 2000-572036/53.
 DR N-PSDB; AAA50340.
 XX
 XX Regulating mucus secretion by a mucus-secreting cell, useful for
 PT treating e.g. bronchitis, asthma or pneumonia, by administering a
 PT compound that inhibits or enhances myristoylated alanine-rich C-kinase
 PT substrate protein -
 XX
 PS Disclosure; Page 46-47; 66pp; English.
 XX The present sequence is that of human myristoylated alanine-rich C
 CC kinase substrate MARCKS protein, a major cellular substrate. The
 CC invention relates to methods of inhibiting mucus secretion by a
 CC mucus-secreting cell by administering a compound that inhibits
 CC MARCKS protein-related mucus secretion. Such compounds include
 CC active fragments of MARCKS protein such as MANS peptide (see
 CC AAY95897) and MA-PDS peptide (see AAY95897), which corresponds to a
 CC phosphorylation site of MARCKS. The inhibitor compounds can be
 CC used to treat conditions such as bronchitis, cystic fibrosis,
 CC chronic obstructive pulmonary disease, asthma, emphysema,
 CC pneumonia, influenza, rhinitis and the common cold. An alternative
 CC sequence for MARCKS is provided in AAY95898, which differs from the
 CC present sequence at 2 amino acid residues, Ser-84 (Ala) and
 CC Ala-119 (Pro).
 XX
 SQ Sequence 332 AA;

Query Match 90.5%; Score 19; DB 21; Length 332;
 Best Local Similarity 66.7%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eagxxs 6
 ||| |
 Db 206 eagaa 211

RESULT 31
 AAY97409

ID AAY97409 standard; Protein; 333 AA.

XX AC AAY97409;

XX XX 14-SEP-2000 (first entry)

XX DE Zebrafish Hsp-binding protein HspBPF.

KW zebrafish; heat shock protein binding protein; HspBPF;
 KW development; apoptosis; cellular stress; heart disease; cancer;
 KW hypoxia.

XX OS Brachydanio rerio.

FH Key Location/Qualifiers

FT Misc-difference 81 /label= unknown

FT /note= "encoded by GAA"

FT Misc-difference 90 /label= unknown

FT /note= "encoded by GAC"

FT Misc-difference 91 /label= unknown

FT /note= "encoded by GAG"

FT WO200031114-A1.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-US27651.

XX 20-NOV-1998; 98US-0109351.

XX (GUER/) GUERRIERO V.

XX (RAYN/) RAYNES D A.

XX Guerriero V, Raynes DA;

XX WPI: 2000-400030/34.

DR N-PSDB; AAA38747.

XX New polynucleotides encoding human heat-shock protein-binding protein,

PT HspBP-1 and HspBP-2, useful for investigating the effects of heat

PT shock-protein regulation -

XX Disclosure; Page 35-36; 38pp; English.

XX The present sequence is the sequence for the zebrafish heat shock

CC protein-binding protein (HspBPF), which is involved in the

CC regulation of the heat shock proteins, which are components of the

CC development, apoptosis and cellular stress pathways. The human

CC homologue is useful not only for research into these areas, but also

CC for treating disorders such as heart disease, hypoxia and cancer.

XX SQ Sequence 333 AA;

Query Match 90.5%; Score 19; DB 21; Length 333;
 Best Local Similarity 66.7%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eagxxs 6
 ||| |
 Db 23 eagas 28

RESULT 32

AAG90789

ID AAG90789 standard; Protein; 341 AA.

XX AC AAG90789;

XX 26-SEP-2001 (first entry)
 XX C glutamicum protein fragment SEQ ID NO: 4543.
 DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis.
 KW Corynebacterium glutamicum.
 KW OS
 XX PN
 XX EP1108790-A2.
 XX 20-JUN-2001.
 XX 18-DEC-2000; 2000EP-0127688.
 XX 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
 PI WPI: 2001-376931/40.
 DR N-PSDB; AAH66008.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX Claim 17; SEQ ID NO: 4543; 246pp + Sequence Listing: English.
 PS The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX SQ Sequence 341 AA;

Query Match 90.5%; Score 19; DB 22; Length 341;
 Best Local Similarity 66.7%; Pred. No. 2.1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 239 eagtss 244

RESULT 33

AA879110
 ID AAB79110 standard; Protein; 341 AA.

XX AAB79110;

AC 30-APR-2001 (first entry)

DT Corynebacterium glutamicum HA protein sequence SEQ ID NO:176.

DE Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
 XX fine chemical production; organic acid; proteinoogenic amino acid;
 KW
 XX

KW nonproteinoogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; genetic engineering;
 KW Brevibacterium; environmental condition.

OS Corynebacterium glutamicum.

XX WO200100842-A2.

PN 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB00911.

XX 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031636.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032127.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032129.

PR 09-JUL-1999; 99DE-1032226.

PR 14-JUL-1999; 99DE-1032920.

PR 14-JUL-1999; 99DE-1032922.

PR 14-JUL-1999; 99DE-1032924.

PR 14-JUL-1999; 99DE-1032928.

PR 14-JUL-1999; 99DE-1032930.

PR 14-JUL-1999; 99DE-1032933.

PR 14-JUL-1999; 99DE-1032935.

PR 14-JUL-1999; 99DE-1032973.

PR 14-JUL-1999; 99DE-1033002.

PR 14-JUL-1999; 99DE-1033003.

PR 14-JUL-1999; 99DE-1033005.

PR 14-JUL-1999; 99DE-1033006.

PR 31-AUG-1999; 99DE-1041378.

PR 31-AUG-1999; 99DE-1041379.

PR 31-AUG-1999; 99DE-1041390.

PR 31-AUG-1999; 99DE-1041391.

PR 03-SEP-1999; 99DE-1042088.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX WPI: 2001-061974/07.

XX N-PSDB; AAF71225.

XX New Isolated Corynebacterium glutamicum nucleic acid for production or
 PT modulation of production of fine chemicals such as amino acids,
 PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
 PT or enzymes -

XX Claim 20; Page 383-384; 712pp; English.

XX AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
 CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The
 CC C. glutamicum HA genes (I) can be used in vectors for expression in host
 CC cells and production of fine chemicals, such as, an organic acid,
 CC proteinoogenic or nonproteinoogenic amino acid (preferred), purine or
 CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
 CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
 CC polyketide or enzyme. The amino acids produced can be lysine, glutamine,
 CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
 CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
 CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
 CC be modulated. The presence of (I) or HA proteins encoded by then are
 CC used for diagnosing the presence or activity of Corynebacterium
 CC diphtheriae. (I) can be used to map the C. glutamicum genome or can be
 CC used as markers for genetically engineered Corynebacterium or
 CC Brevibacterium. The HA proteins encoded by the (I) are used to maintain
 CC homeostasis in C. glutamicum or help the microorganism to adapt to
 CC different environmental conditions.

SQ Sequence 341 AA;
Query Match 90.5%; Score 19; DB 22; Length 341;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
Db 239 eagtss 244
RESULT 34
AAB79144
ID AAB79144 standard; Protein; 341 AA.
AC AAB79144;
XX
XX 30-APR-2001 (first entry)
DE Corynebacterium glutamicum HA protein sequence SEQ ID NO:244.
DE Corynebacterium glutamicum; homeostasis; adaptation: HA protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; genetic engineering;
KW Brevibacterium; environmental condition.
XX
XX Corynebacterium glutamicum.
OS
XX WO200100842-A2.
PN
XX 04-JAN-2001.
PD
XX
XX 23-JUN-2000; 2000WO-IB00911.
PF
XX 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032127.
PR 09-JUL-1999; 99DE-1032128.
PR 09-JUL-1999; 99DE-1032129.
PR 09-JUL-1999; 99DE-1032226.
PR 14-JUL-1999; 99DE-1032920.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1032930.
PR 14-JUL-1999; 99DE-1032933.
PR 14-JUL-1999; 99DE-1032935.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033002.
PR 14-JUL-1999; 99DE-1033003.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041390.
PR 31-AUG-1999; 99DE-1041391.
PR 03-SEP-1999; 99DE-1042088.
XX
XX (BADI) BASF AG.
PA
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI WPI; 2001-061974/07.
XX N-PSDB; AAF71259.
DR
XX New isolated Corynebacterium glutamicum nucleic acid for production or
PT modulation of production of fine chemicals such as amino acids,

PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
or enzymes -
XX
PS Claim 20; Page 466-467; 712pp; English.
XX
CC AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
and adaptation (HA) proteins given in AAB79023 to AAB79242. The
C. glutamicum HA genes (I) can be used in vectors for expression in host
cells and production of fine chemicals, such as, an organic acid,
proteinogenic or nonproteinogenic amino acid (preferred), purine or
pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
polyketide or enzyme. The amino acids produced can be lysine, glutamine,
glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
tyrosine, phenylalanine, or tryptophan. The fine chemical production can
be modulated. The presence of (I) or HA proteins encoded by then are
used for diagnosing the presence or activity of Corynebacterium
diphtheriae. (I) can be used to map the C. glutamicum genome or can be
used as markers for genetically engineered Corynebacterium or
Brevibacterium. The HA proteins encoded by the (I) are used to maintain
homeostasis in C. glutamicum or help the microorganism to adapt to
different environmental conditions.
XX
XX Sequence 341 AA;
SQ
Query Match 90.5%; Score 19; DB 22; Length 341;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
Db 239 eagtss 244
RESULT 35
AAW44368
ID AAW44368 standard; Protein; 349 AA.
XX
AC AAW44368;
XX
DT 29-MAY-1998 (first entry)
XX
DE Aspergillus nidulans metallo-protease pepH.
XX
KW Aspergillus nidulans; metallo-protease; pepH; protein degradation;
KW fungus; food processing.
XX
OS Aspergillus nidulans.
XX
PN WO9746689-A1.
XX
PD 11-DEC-1997.
XX
PF 05-JUN-1997; 97WO-EF02982.
XX
PR 05-JUN-1996; 96EP-0201579.
XX
PA (KONN) GIST-BROCADES BV.
XX
PI Van Den Hombergh JPTW, Visser J;
XX
DR WPI; 1998-042197/04.
DR N-PSDB; AAV15305, AAV15306.
XX
PT Metallo-protease deficient fungus with site selected DNA disruption
- and Aspergillus metallo-protease genes, useful in protein
production to reduce protease activity hence protein degradation
XX
PS Example 4; Page 26-27; 53pp; English.
XX
CC The present sequence represents a metallo-protease, pepH, from

SQ Sequence 363 AA;

Query Match 90.5%; Score 19; DB 22; Length 363;
Best Local Similarity 66.7%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eagxxs 6
||| |
Db 196 eagass 201

RESULT 38

AAY05663
ID AAY05663 standard; Protein; 366 AA.

XX AAY05663;

XX 19-JUL-1999 (first entry)

XX Maize caffeic O-methyltransferase.

XX Maize; corn; caffeic O-methyltransferase; lignin; transgenic plant.

XX Zea mays.

XX WO9910498-A2.

XX 04-MAR-1999.

XX 24-AUG-1998; 98WO-US17519.

XX 12-MAY-1998; 98US-0076851.

XX 27-AUG-1997; 97US-0057082.

XX (PION-) PIONEER HI-BRED INT INC.

XX Bowen BA, Helentjaris TG, Wang X;

XX WPI; 1999-204667/17.

XX N-PSDB; AAX25202.

PT Nucleic acids encoding plant lignin biosynthesis enzymes - used to
transform plants to modulate lignin biosynthesis

XX Claim 9; Page 96-97; 166pp; English.

XX The present sequence is a caffeic O-methyltransferase of maize,
CC encoded by a clone (see AAX25202) isolated from a maize line B73 cDNA
CC library. The invention provides methods and compositions relating
CC to altering lignin biosynthesis and/or the lignin composition of
CC plants. Isolated nucleic acids (see AAX25196-216) that code for
CC proteins (see AAY05657-77) involved in lignin biosynthesis are
CC claimed. Also claimed are recombinant expression cassettes, host
CC cells (especially maize or sorghum), and transgenic plants and
CC seeds. The claimed nucleic acids can be used to transform a plant
CC to modulate lignin biosynthesis. A claimed method involves
CC transforming a plant cell with a recombinant expression cassette
CC comprising a lignin biosynthesis polynucleotide operably linked to
CC a promoter, growing the plant cell under plant growing conditions,
CC and inducing expression of the polynucleotide for a time sufficient
CC to modulate (preferably increase) lignin biosynthesis in the plant.
CC The plant lignins can be used as chemical feedstock. plant
CC material of increased lignin content can be used as a fuel source,
CC and in the pulp and paper industry. Decreased lignin content
CC improves the digestibility of fodder crops.

XX Sequence 366 AA;

Query Match 90.5%; Score 19; DB 20; Length 366;
Best Local Similarity 66.7%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eagxxs 6
||| |
Db 115 eagtas 120

RESULT 39

AAG20561

ID AAG20561 standard; Protein; 377 AA.

XX AAG20561;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 22802.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 18-MAY-1999; 99US-0134370.

XX 19-MAY-1999; 99US-0134768.

XX 20-MAY-1999; 99US-0134941.

XX 21-MAY-1999; 99US-0135124.

XX 24-MAY-1999; 99US-0135353.

XX 25-MAY-1999; 99US-0135629.

XX 27-MAY-1999; 99US-0136021.

XX 28-MAY-1999; 99US-0136392.

XX 01-JUN-1999; 99US-0136782.

XX 03-JUN-1999; 99US-0137222.

XX 04-JUN-1999; 99US-0137528.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 14-JUN-1999; 99US-0138847.

XX 16-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 377;
Best Local Similarity 66.7%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY 1 eagxxx 6
||| |
Db 75 eagsss 80

RESULT 40

AAG41800
ID AAG41800 standard; Protein; 377 AA.
AC AAG41800;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52052.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 377;
Best Local Similarity 66.7%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaaxxs 6
||| |
Db 75 eagsss 80

RESULT 41

ID AAG24042
AC AAG24042 standard; Protein; 386 AA.

XX AAG24042;

XX AAG24042;

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27567.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

```
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
```

```
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 90.5%; Score 19; DB 21; Length 386;
Best Local Similarity 66.7%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 46 eaqaas 51

RESULT 42
AAP70581
ID AAP70581 standard; Protein; 390 AA.
XX
AC AAP70581;
XX
DT 03-MAY-1991 (first entry)
XX
DE Protease biosynthetic protein.
XX
KW Saccharomycosis; yeast.
XX
PN JP62104578-A.
XX
PD 15-MAY-1987.

XX 31-OCT-1985; 85JP-0244893.
 XX 31-OCT-1985; 85JP-0244893.
 XX (FUKU/) FUKUI S.
 XX WPI; 1987-173695/25.
 DR N-PSDB; AAN70927.
 XX
 PT Protease prodn. - by culturing microorganism transformed with
 PT vector derived from saccharomycosis.
 XX
 PS Disclosure; Fig 1; 9pp; Japanese.
 XX
 CC Product is a biosynthetic component involved in the synthesis of
 CC protease. The protein may be produced from a transformed *S.cerevisiae*
 CC expression system for the large scale production of protease.
 XX
 SQ Sequence 390 AA;

Query Match 90.5%; Score 19; DB 8; Length 390;
 Best Local Similarity 66.7%; Pred. No. 2.4e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxss 6
 Db 266 eagss 271

RESULT 43
 AAY35147
 ID AAY35147 standard; Protein; 393 AA.
 XX AC AAY35147;
 XX
 DT 13-SEP-1999 (first entry)
 XX
 DE Chlamydia pneumoniae transmembrane protein sequence.
 XX
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.
 XX
 XX Chlamydia pneumoniae.
 OS WO9927105-A2.
 XX
 PN 03-JUN-1999.
 XX
 XX 20-NOV-1998; 98WO-IB01890.
 XX
 XX 04-NOV-1998; 98US-0107078.
 PR 21-NOV-1997; 97FR-0014673.
 XX
 XX (GEST) GENSET.
 PA
 XX Griffais R;
 PI
 XX WPI; 1999-357842/30.
 DR
 XX
 XX Genome sequence of Chlamydia pneumoniae
 PT
 PS Page 1016; Disclosure; 1912pp; English.
 XX
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in

CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.
 XX
 SQ Sequence 393 AA;

Query Match 90.5%; Score 19; DB 20; Length 393;
 Best Local Similarity 66.7%; Pred. No. 2.4e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxss 6
 Db 101 eagss 106

RESULT 44
 AAY28643
 ID AAY28643 standard; Protein; 422 AA.
 XX AC AAY28643;
 XX
 DT 03-NOV-1999 (first entry)
 XX
 DE Human serine protease inhibitor from cDNA clone HETDK50.
 XX
 KW Human serine protease inhibitor from cDNA clone HETDK50; fusion protein;
 KW serpin; serine protease; human pre-alpha-1-antitrypsin precursor;
 KW extracellular matrix degradation; multiple sclerosis; cancer; arthritis;
 KW inflammation; immune system disorder; neurodegenerative disorder;
 KW kallmann's syndrome; Down's syndrome; Alzheimer's; secreted protein;
 KW galactorrhea; hypogonadism; somatostatin; protein purification.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Peptide I..19
 FT /label= Signal_peptide
 FT 20..422
 FT /label= Mature_serine_protease_inhibitor

WO9940183-A1.
 XX
 XX 12-AUG-1999.
 XX
 XX 04-FEB-1999; 99WO-US02292.
 XX
 XX 06-FEB-1998; 98US-0073961.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ni J, Ruben SM;
 PI
 XX WPI; 1999-508502/42.
 DR N-PSDB; AAX80907.
 XX
 XX New isolated human serine protease and serpin polypeptides, used to
 PT develop products for treating e.g. immune disorders, cancers,
 PT inflammation, transplant rejection or infections, or as food
 PT additives
 PT
 XX Claim 11; Pages 83-85; 99pp; English.
 PS
 XX The present sequence is a serine protease inhibitor (serpin) from cDNA
 CC clone HETDK50 which is obtained from human endometrial tumour tissue
 CC cDNA library. The protein shows a high degree of sequence similarity to
 CC human pre-alpha-1-antitrypsin precursor. The serpin and its
 CC coding sequence are used in the diagnosis and treatment of disorders
 CC related to abnormal level of the protein or mutation in the nucleotide
 CC sequence. The serpin can be used for treating disorders characterised by
 CC degradation of extracellular matrix, e.g. cancer, arthritis, multiple
 CC sclerosis and immune system disorders, for treating wasting associated

CC with excessive protease production during inflammation or
 CC neurodegenerative disorders e.g. Kallmann's and Down's syndromes,
 CC Alzheimer's and Huntington's diseases. It may also be used to reduce
 CC excess levels of prolactin in the treatment of galactorrhea and
 CC hypogonadism, and decrease the amount of free circulating somatostatin to
 CC prevent somatostatin's inhibitory effect on the release of growth
 CC hormone. The fusion of this protein to His-tag, HA-tag, Igg domains,
 CC etc. facilitates protein purification and fusion to Igg-1, Igg-3 and
 CC albumin increases the half life time in vivo.
 XX
 SQ Sequence 422 AA;

Query Match 90.5%; Score 19; DB 20; Length 422;
 Best Local Similarity 66.7%; Pred. No. 2.6e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||| |
 Db 370 eagaas 375

RESULT 45
 AAB74691
 ID AAB74691 standard; Protein; 422 AA.

AC AAB74691;

XX 12-JUN-2001 (first entry)

XX Human protease and protease inhibitor PPIM-24.

XX Human; protease; protease inhibitor; protease and protease inhibitor;
 KW PPIM; identification; diagnosis; anti-human immunodeficiency virus; HIV;
 KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
 KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
 KW cytostatic; antibacterial; fungicide; protozoacide; antiarteriosclerotic;
 KW antiatherosclerotic; antiapoptotic; virucide; hepatotropic; gene therapy;
 KW autoimmune disorder; inflammatory disease; SCID; Chediak-Higashi syndrome;
 KW severe combined immunodeficiency disease; Addison's disease; autoimmune thyroiditis; gout;
 KW Cushing's disease; Hashimoto's thyroiditis; Sjogren's syndrome; infection;
 KW Crohn's disease; diabetes mellitus; Good pasture's syndrome; cancer;
 KW Grave's disease; Hashimoto's thyroiditis; Sjogren's syndrome; cancer;
 KW Werner's syndrome; cell proliferative disorder; arteriosclerosis;
 KW atherosclerosis; cirrhosis; hepatitis; psoriasis.

XX Homo sapiens.

XX WO200110903-A2.

XX 15-FEB-2001.

XX 09-AUG-2000; 2000WO-US21878.

XX 09-AUG-1999; 99US-0147986.

XX 21-OCT-1999; 99US-0160807.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lal P, Tang YT, Bandman O, Baughn MR, Azimzai Y, Lu DM;

XX Yang J;

XX WPI; 2001-202760/20.

XX N-PSDB; AAF81737.

XX New protease (inhibitors) useful for diagnosis and treatment of
 PT autoimmune/inflammatory disorders such as acquired immunodeficiency
 PT syndrome, Cushing's disease, Addison's disease and cell proliferative
 PT disorders such as cancer -

XX Claim 1; Page 112-113; 134pp; English.

XX AAF81714 to AAF81740 encode the human proteases and protease inhibitors

CC (PPIMs) given in AAB74668 to AAB74694. The PPIMs can have activities such
 CC as: anti-human immunodeficiency virus (HIV); antidiabetic; antithyroid;
 CC immunostimulant; immunomodulator; antiinflammatory; immunosuppressive;
 CC nephrotropic; antigout; thyromimetic; cytostatic; antibacterial;
 CC fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic;
 CC virucide; antipsoriatic; and hepatotropic. PPIM polynucleotide and
 CC protein sequences can be used in the diagnosis, treatment and prevention
 CC of autoimmune/inflammatory disorders such as AIDS, DiGeorge's syndrome,
 CC severe combined immunodeficiency disease (SCID), Chediak-Higashi
 CC syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,
 CC Crohn's disease, diabetes mellitus, Good pasture's syndrome, gout,
 CC Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's
 CC syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic
 CC infections and cell proliferative disorder such as arteriosclerosis,
 CC atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPIM
 CC polynucleotide sequences can be used in somatic or germline gene therapy
 CC and in diagnosis of diseases. They can also be used in generating
 CC hybridisation probes useful in mapping the naturally occurring genomic
 XX sequences and in molecular biology techniques.

SQ Sequence 422 AA;

Query Match 90.5%; Score 19; DB 22; Length 422;
 Best Local Similarity 66.7%; Pred. No. 2.6e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||| |
 Db 370 eagaas 375

RESULT 46

AAG41388

ID AAG41388 standard; Protein; 427 AA.

XX AAG41388;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 51488.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0129845.

XX 21-APR-1999; 99US-0130077.

XX 23-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 427;
Best Local Similarity 66.7%; Pred.No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
   ||| |
Db 6 eagsss 11
      .

RESULT 47
AAG41387
ID AAG41387 standard; Protein; 431 AA.
AC AAG41387;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51487.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
```

PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 08-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 431;
Best Local Similarity 66.7%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
| | | |
Db 10 eagsss 15

RESULT 48
AAG24041
ID AAG24041 standard; Protein; 435 AA.
XX AC AAG24041;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 27566.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 22-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154019.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```
PR 29-OCT-1999; 99US-0162142.
Query Match 90.5%; Score 19; DB 21; Length 435;
Best Local Similarity 66.7%; Pred. NO. 2.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
   ||| |
Db 95 eagaas 100

RESULT 49
AAG24040
ID AAG24040 standard; Protein; 440 AA.
XX AC AAG24040;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 27565.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX QY 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
99US-0138540.
99US-0138647.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144335.
99US-0144352.
99US-0144632.
99US-0144884.
99US-0144814.
99US-0145086.
99US-0145088.
99US-0145085.
99US-0145087.
99US-0145089.
99US-0145192.
99US-0145145.
99US-0145218.
99US-0145224.
99US-0145276.
99US-0145913.
99US-0145918.
99US-0145919.
99US-0145951.
99US-0146386.
99US-0146388.
99US-0146389.
99US-0147038.
99US-0147204.
99US-0147302.
99US-0147192.
99US-0147260.
99US-0147303.
99US-0147416.
99US-0147493.
99US-0147493.
```

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 440;
Best Local Similarity 66.7%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6

Db 100 eagaaas 105
||| |

RESULT 50

AAR31955

ID AAR31955 standard; Protein; 441 AA.

XX AAR31955;

XX AAR31955;

DT 06-JUN-1993 (first entry)

XX

DE Sequence encoded by glycoprotein G gene.

XX

KW IBR glycoprotein E gene; unique short 2 gene.

XX

OS Infectious bovine rhinotracheitis.

XX

PN WO9302104-A.

XX

PD 04-FEB-1993.

XX

PF 20-JUL-1992; 92WO-US06034.

XX

PR 18-JUL-1991; 91US-0732584.

XX

(SYTR) SYNTRO CORP.

XX

PI Cochran MD, Macdonald RD;

XX

DR WPI; 1993-058725/07.

XX

DR N-PSDB; AAQ36768.

XX

PT Recombinant infectious bovine rhinotracheitis virus - provides

PT isolated DNA encoding gpE glyco:protein, gpG glyco:protein and

PT unique short 2 genes of the virus

XX

PS Example; Fig 8; 240pp; English.

XX

CC The sequence of approximately 1400 base pairs of the HindIII K

CC fragment, starting approximately 2800 base pairs downstream of the

CC HindIII K/HindIII O junction, are shown. The glycoprotein G (gpG)

CC gene is transcribed away from the HindIII K/HindIII O junction.

XX

SQ Sequence 441 AA;

Query Match 90.5%; Score 19; DB 14; Length 441;

Best Local Similarity 66.7%; Pred. No. 2.7e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6

||| |

Db 265 eagsas 270

Search completed: September 24, 2002, 11:22:18
Job time: 6017 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 10:08:51 ; Search time 36.48 Seconds
(without alignments)
4.017 Million cell updates/sec

Title: BASK-853-CLAIM4

Perfect score: 21

Sequence: 1 eagxs 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

Issued Patents.AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	90.5	47	4	Sequence 345, App
2	19	90.5	159	4	Sequence 2, Appli
3	19	90.5	211	4	Sequence 4, Appli
4	19	90.5	332	2	Sequence 5, Appli
5	19	90.5	441	1	Sequence 21, Appl
6	19	90.5	441	2	Sequence 21, Appl
7	19	90.5	512	4	Sequence 6, Appli
8	19	90.5	1088	5	Sequence 3, Appli
9	19	90.5	1183	2	Sequence 2, Appli
10	19	90.5	9	4	Sequence 2, Appli
11	18	85.7	59	2	Sequence 31, Appl
12	18	85.7	59	2	Sequence 109, App
13	18	85.7	59	4	Sequence 109, App
14	18	85.7	59	5	Sequence 109, App
15	18	85.7	61	4	Sequence 12, Appl
16	18	85.7	66	4	Sequence 10, Appl
17	18	85.7	91	2	Sequence 19, Appl
18	18	85.7	91	4	Sequence 19, Appl
19	18	85.7	123	1	Sequence 2, Appli
20	18	85.7	175	4	Sequence 9, Appli
21	18	85.7	211	2	Sequence 2, Appli
22	18	85.7	428	4	Sequence 2, Appli
23	18	85.7	433	2	Sequence 1, Appli
24	18	85.7	481	3	Sequence 1, Appli
25	18	85.7	500	4	Sequence 25, Appl
26	18	85.7	503	1	Sequence 2, Appli
27	18	85.7	503	4	Sequence 52, Appl

US-08-936-165A-345
Sequence 345, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

ALIGNMENTS

RESULT 1.

US-08-936-165A-345
Sequence 345, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 345:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-345

Query Match 90.5%; Score 19; DB 4; Length 47;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
||| |

Db 31 EAGATS 36

RESULT 2
US-08-856-253-2
; Sequence 2, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856.253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-2

Query Match 90.5%; Score 19; DB 4; Length 159;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
||| |

Db 28 EAGTSS 33

RESULT 3
US-08-856-253-4
; Sequence 4, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856.253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-4

Query Match 90.5%; Score 19; DB 4; Length 211;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
||| |

Db 34 EAGTSS 39

RESULT 4
US-08-405-175A-5
; Sequence 5, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue

CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,175A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-121A
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Predicted primary structure of human MARCKS
HYPOTHETICAL: NO
US-08-405-175A-5

Query Match 90.5%; Score 19; DB 2; Length 332;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 206 EAGAS 211

RESULT 5
US-08-191-866D-21
Sequence 21, Application US/08191866D
Patent No. 5783195
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Macdonald, Richard D.
TITLE OF INVENTION: Recombinant Infectious Bovine
TITLE OF INVENTION: Rhinotracheitis Virus S-IBR-052 And Uses Thereof
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191,866D
FILING DATE: 4 February 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-191-866D-21

Query Match 90.5%; Score 19; DB 1; Length 441;
Best Local Similarity 66.7%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 265 EAGSAS 270

RESULT 6
US-08-185-949B-21
Sequence 21, Application US/08185949B
Patent No. 5874279
GENERAL INFORMATION:
APPLICANT: Mark D. Cochran
APPLICANT: Richard D. Macdonald
TITLE OF INVENTION: Recombinant Infectious Bovine
TITLE OF INVENTION: Rhinotracheitis Virus
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,949B
FILING DATE: 03-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 278-0525
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-185-949B-21

Query Match 90.5%; Score 19; DB 2; Length 441;
Best Local Similarity 66.7%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 265 EAGSAS 270

RESULT 7
US-08-856-253-6
Sequence 6, Application US/08856253

; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-856-253-6

Query Match 90.5%; Score 19; DB 4; Length 512;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 149 EAGTSS 154

RESULT 8
PCT-US95-03747-3
; Sequence 3, Application PC/TUS9503747
; GENERAL INFORMATION:
; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
; TITLE OF INVENTION: Brevican, A Glial Cell Proteoglycan
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03747
; FILING DATE: 27-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Imbra, Richard J.
; REGISTRATION NUMBER: 37,643
; REFERENCE/DOCKET NUMBER: FP-LJ 1453
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 908 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US95-03747-3

Query Match 90.5%; Score 19; DB 5; Length 908;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 583 EAGSSS 588

RESULT 9
US-09-130-242-2
; Sequence 2, Application US/09130242B
; Patent No. 6194558
; GENERAL INFORMATION:
; APPLICANT: Gianturco, S.H.
; APPLICANT: Bradley, W.A.
; TITLE OF INVENTION: DNA Encoding Human Monocyte-Macrophage Aoplipoprotein
; FILE REFERENCE: D5880
; CURRENT APPLICATION NUMBER: US/09/130,242B
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: MS WORD, Macintosh OS 8.5
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-130-242-2

Query Match 90.5%; Score 19; DB 4; Length 1088;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 327 EAGTAS 332

RESULT 10
US-08-447-031A-2
; Sequence 2, Application US/08447031A
; Patent No. 5851794
; GENERAL INFORMATION:
; APPLICANT: GUSS, Bengt
; APPLICANT: HOOK, Magnus
; APPLICANT: JONSSON, Hans
; APPLICANT: LINDBERG, Martin
; APPLICANT: PATTI, Joseph
; APPLICANT: SWITALSKI, Lech
; TITLE OF INVENTION: A COLLAGEN BINDING PROTEIN AS WELL AS
; NUMBER OF SEQUENCES: 8

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/447,031A
;; FILING DATE: 22-MAY-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/861,804
;; FILING DATE: 21-AUG-1992
;; PRIOR APPLICATION DATA: WO PCT/SE91/00707
;; APPLICATION NUMBER: SE 9003374-7
;; FILING DATE: 22-OCT-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 012889-006
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1183 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-447-031A-2

Query Match 90.5%; Score 19; DB 2; Length 1183;
Best Local Similarity 66.7%; Fred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
Db 166 EAGTSS 171

RESULT 11
US-08-485-355B-31
; Sequence 31, Application US/08485355B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H. J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B

;;
;; FILING DATE: 07-Jun-1995
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/440,522
;; FILING DATE: 12-MAY-1995
;; APPLICATION NUMBER: US 08/089,372
;; FILING DATE: 08-JUL-1993
;; APPLICATION NUMBER: AU PL4081/92
;; FILING DATE: 14-AUG-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trecartin, Richard F.
;; REGISTRATION NUMBER: 31,801
;; REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 398-3249
;; TELEX: 910 277299
;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-08-485-355B-31

Query Match 85.7%; Score 18; DB 4; Length 9;
Best Local Similarity 66.7%; Fred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
Db 3 EAGVAS 8

RESULT 12
US-08-596-387B-109
; Sequence 109, Application US/08596387B
; Patent No. 5869270
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Jiao, Jin-An
; APPLICANT: Burkhardt, Martin
; APPLICANT: Wong, Hing
; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,387B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09816
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,302
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:

NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-596-387B-109

Query Match 85.7%; Score 18; DB 2; Length 59;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 41 EAGRAS 46

RESULT 13
US-09-067-615-109
Sequence 109, Application US/09067615
Patent No. 6309645
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Jiao, Jin-An
APPLICANT: Burkhardt, Martin
APPLICANT: Wong, Hing C.
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,615
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/596,387
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

US-09-067-615-109

Query Match 85.7%; Score 18; DB 4; Length 59;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 41 EAGRAS 46

RESULT 14
PCT-US95-09816A-109
Sequence 109, Application PC/TUS9509816A
GENERAL INFORMATION:
APPLICANT: Wong, Hing C.
APPLICANT: Rhode, Peter R.
APPLICANT: Widanz, Jon A.
APPLICANT: Grammer, Susan
APPLICANT: Edwards, Ana C.
APPLICANT: Chavallaz, Pierre-Andre
APPLICANT: Jiao, Jin-An
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09816A
FILING DATE: 31-JUL-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
PCT-US95-09816A-109

Query Match 85.7%; Score 18; DB 5; Length 59;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 41 EAGRAS 46

RESULT 15

US-08-446-137B-12
; Sequence 12, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trower, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; TITLE OF INVENTION: FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-446-137B-12

Query Match 85.7%; Score 18; DB 4; Length 61;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 29 EAGITS 34

RESULT 16
US-08-446-137B-10
; Sequence 10, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trower, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; TITLE OF INVENTION: FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

US-08-446-137B-10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-446-137B-10

Query Match 85.7%; Score 18; DB 4; Length 66;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 33 EAGISS 38

RESULT 17
US-08-598-873-19
; Sequence 19, Application US/08598873
; Patent No. 5928884
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo M.
; APPLICANT: Huebner, Kay
; TITLE OF INVENTION: PHIT PROTEINS AND NUCLEIC ACIDS AND
; TITLE OF INVENTION: METHODS BASED THEREON
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/598,873
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Friebe, Thomas E.
; REGISTRATION NUMBER: 29,258
; REFERENCE/DOCKET NUMBER: 8666-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown

```
; MOLECULE TYPE: peptide
; US-08-598-873-19

Query Match      85.7%; Score 18; DB 2; Length 91;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 49 EAGKSS 54

RESULT 18
US-08-605-430-19
; Sequence 19, Application US/08605430
; Patent No. 6242212
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo M.
; APPLICANT: Huebner, Kay
; TITLE OF INVENTION: PHIT PROTEINS AND NUCLEIC ACIDS AND
; METHODS BASED THEREON
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,430
; FILING DATE: 22-FEB-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Friebe, Thomas E.
; REGISTRATION NUMBER: 29,258
; REFERENCE/DOCKET NUMBER: 8666-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-605-430-19

Query Match      85.7%; Score 18; DB 4; Length 91;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 49 EAGKSS 54

RESULT 19
US-07-869-912-2
; Sequence 2, Application US/07869912
; Patent No. 5316922
; GENERAL INFORMATION:
; APPLICANT: Court, Don
; APPLICANT: Brown, Stanley
; TITLE OF INVENTION: A Method for Identifying and
```

```
; TITLE OF INVENTION: Expressing Proteins that Recognize and Adhere to Specific
; TITLE OF INVENTION: Probes
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth A. Weber
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,912
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 32,334
; REFERENCE/DOCKET NUMBER: 15280-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-869-912-2

Query Match      85.7%; Score 18; DB 1; Length 123;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 21 EAGGSS 26

RESULT 20
US-08-446-137B-9
; Sequence 9, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
```

ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 100084.406
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 175 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-446-137B-9

Query Match 85.7%; Score 18; DB 4; Length 175;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
DB 29 EAGTGS 34

RESULT 21
US-08-708-958-2
Sequence 2, Application US/08708958
Patent No. 5948952
GENERAL INFORMATION:
APPLICANT: SANDS, Arthur T.
APPLICANT: BRADLEY, Allan
APPLICANT: ABUIN, Alejandro
TITLE OF INVENTION: XERODERMA PIGMENTOSUM-DEFICIENT
TITLE OF INVENTION: MOUSE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/708,958
FILING DATE: SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6641
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-958-2

Query Match 85.7%; Score 18; DB 2; Length 211;
Best Local Similarity 66.7%; Pred. No. 7.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
DB 25 EAGSGS 30

RESULT 22
US-09-423-340-2
Sequence 2, Application US/09423340
Patent No. 6225454
GENERAL INFORMATION:
APPLICANT: MIYAGI, Taeko
APPLICANT: WADA, Tadashi
APPLICANT: YOSHIKAWA, Yuko
TITLE OF INVENTION: SIALIDASE LOCALIZED IN PLASMA MEMBRANE AND
FILE REFERENCE: OP699
CURRENT APPLICATION NUMBER: US/09/423,340
CURRENT FILING DATE: 1999-11-22
EARLIER APPLICATION NUMBER: JP 9-132174
EARLIER FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 428
TYPE: PRT
ORGANISM: Bos primigenius
US-09-423-340-2

Query Match 85.7%; Score 18; DB 4; Length 428;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
DB 321 EAGTLS 326

RESULT 23
US-08-749-902-1
Sequence 1, Application US/08749902
Patent No. 5985635
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
TITLE OF INVENTION: PROTEIN KINASES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,902
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0150 US
TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: Consensus
US-08-749-902-1

Query Match 85.7%; Score 18; DB 2; Length 433;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxys 6
||| |
Db 396 EAGAVS 401

RESULT 24
US-09-330-095-1
Sequence 1, Application US/09330095
Patent No. 6127161
GENERAL INFORMATION:
APPLICANT: Kikkoman Corporation
TITLE OF INVENTION: Leucine Aminopeptidase Gene, Recombinant DNA, and
FILE REFERENCE: PH-622
CURRENT APPLICATION NUMBER: US/09/330,095
CURRENT FILING DATE: 1999-06-11
EARLIER APPLICATION NUMBER: JP-164611/1998
EARLIER FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 481
TYPE: PRT
ORGANISM: Aspergillus sojae
US-09-330-095-1

Query Match 85.7%; Score 18; DB 3; Length 481;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxys 6
||| |
Db 193 EAGSVS 198

RESULT 25
US-08-960-190A-25
Sequence 25, Application US/08960190A
Patent No. 6232445
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Acevedo, Jorge
APPLICANT: Burkhardt, Martin
APPLICANT: Jiao, Jin-an
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: usa
ZIP: 02109
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,190A
FILING DATE: 29-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 500 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-960-190A-25

Query Match 85.7%; Score 18; DB 4; Length 500;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxys 6
||| |
Db 41 EAGRAS 46

RESULT 26
US-07-612-673-2
Sequence 2, Application US/07612673
Patent No. 5260208
GENERAL INFORMATION:
APPLICANT: Petre, Dominique
APPLICANT: Cerbelaud, Edith
APPLICANT: Mayaux, Jean-Francois
APPLICANT: Yeh, Patrice
TITLE OF INVENTION: NOVEL POLYPEPTIDES, THE DNA SEQUENCES
ALLOWING THEIR EXPRESSION, METHOD OF PREPARATION, AND
UTILIZATION
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/612,673
FILING DATE: 19901114
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 03715.0010
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-612-673-2

Query Match 85.7%; Score 18; DB 1; Length 503;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 149 EAGGS 154

RESULT 27

US-08-845-258-52
Sequence 52, Application US/08845258
Patent No. 6183976

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
APPLICANT: Sleath, Paul R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
AND TREATMENT OF B. MICROTI INFECTION
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/845,258
FILING DATE: 24-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.

REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.426C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 52:

SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Babesia Microti

US-08-845-258-52

Query Match 85.7%; Score 18; DB 4; Length 503;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 417 EAGGTS 422

RESULT 28

US-08-990-571-52
Sequence 52, Application US/08990571
Patent No. 6214971

GENERAL INFORMATION:

APPLICANT: Reed, Steven G. et al.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B.
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/990,571
FILING DATE: 11-DEC-1997

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.426C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 52:

SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Babesia Microti

US-08-990-571-52

Query Match 85.7%; Score 18; DB 4; Length 503;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 417 EAGGTS 422

RESULT 29

US-07-796-361A-11

Sequence 11, Application US/07796361A
Patent No. 5258292

GENERAL INFORMATION:

APPLICANT: YEH, Petrice

APPLICANT: MAYAUX, Jean-Francois

APPLICANT: CERBELAUD, Edith

APPLICANT: PETRE, Dominique

TITLE OF INVENTION: ENZYMIC PROCESS FOR THE SYNTHESIS OF

TITLE OF INVENTION: AMMONIUM ADIPATE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker and Mathis

STREET: The George Mason Building, Washington &

STREET: Prince Streets

CITY: Alexandria

STATE: Virginia

; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/796,361A
; FILING DATE: 19911122
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90-14 853
; FILING DATE: 28-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: CRANE-FEURY, SHARON E.
; REGISTRATION NUMBER: F36,113
; REFERENCE/DOCKET NUMBER: 003025-010
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; TELEX: 440580
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-796-361A-11

Query Match 85.7%; Score 18; DB 1; Length 521;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqx5 6
||| |
Db 167 EAGGSS 172

RESULT 30
US-08-539-666-2
; Sequence 2, Application US/08539666
; Patent No. 5766918
; GENERAL INFORMATION:
; APPLICANT: Petre, Dominique
; APPLICANT: Cerbelaud, Edith
; APPLICANT: Mayaux, Jean-Francois
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: No. 5766918el Polypeptides, The DNA Sequences
; TITLE OF INVENTION: Allowing their Expression, Method of Preparation, and
; TITLE OF INVENTION: Utilization
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,666
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/097,009
; FILING DATE: 27-JUL-1993

; APPLICATION NUMBER: US 07/612,673
; FILING DATE: 14-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 8916332
; FILING DATE: 11-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03715.0010-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-539-666-2

Query Match 85.7%; Score 18; DB 1; Length 521;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqx5 6
||| |
Db 167 EAGGSS 172

RESULT 31
US-08-348-891A-2
; Sequence 2, Application US/08348891A
; Patent No. 5654136
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Keiko
; APPLICANT: MORI, Takayuki
; APPLICANT: MAKINO, Satoshi
; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE,
; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
; TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,891A
; FILING DATE: 25-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,400
; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-293625
; FILING DATE: 14-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: KP-7501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-521-2297
; TELEFAX: 703-685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-348-891A-2

Query Match 85.7%; Score 18; DB 1; Length 525;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxss 6
||| |
Db 263 EAGLAS 268

RESULT 32
US-08-905-817-2 : Sequence 2, Application US/08905817
Patent No. 5824777
GENERAL INFORMATION:
APPLICANT: SASAKI, Keiko
APPLICANT: MORI, Takayuki
APPLICANT: MAKINO, Satoshi
TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE.
TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,817
FILING DATE: 04-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,891
FILING DATE: 25-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,400
FILING DATE: 10-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-293625
FILING DATE: 14-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: KP 7501A
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-905-817-2

Query Match 85.7%; Score 18; DB 2; Length 525;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxss 6
||| |
Db 263 EAGLAS 268

RESULT 33
US-08-513-841-1 : Sequence 1, Application US/08513841
Patent No. 5753481
GENERAL INFORMATION:
APPLICANT: Niwa, Mineo
APPLICANT: Saito, Yoshimasa
APPLICANT: Ishii, Yoshinori
APPLICANT: Yoshida, Masaru
APPLICANT: Suzuki, Hiromi
TITLE OF INVENTION: No. 5753481el L-sorbose Dehydrogenase and No. 5753481el L-s
TITLE OF INVENTION: Dehydrogenase Obtained from Gluconobacter oxydans T-100
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
STREET: 1755 Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS Editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,841
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9304700.9
FILING DATE: 08-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 241851/1993
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-909-0 PCT
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Gluconobacter oxydans
STRAIN: T-100
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..530
IDENTIFICATION METHOD: experimentally
US-08-513-841-1

Query Match 85.7%; Score 18; DB 1; Length 530;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxss 6
||| |
Db 364 EAGVTS 369

```
RESULT 34
US-08-696-834-1
; Sequence 1, Application US/08696834
; Patent No. 5834263
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mineo
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshinori
; APPLICANT: Yoshida, Masaru
; APPLICANT: Hayashi, Hiromi
; TITLE OF INVENTION: Method for Producing 2-Keto-L-Gulonic Acid
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt,
; STREET: 1755 Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,834
; FILING DATE: 24-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 28612/1994
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Gluconobacter oxydans
; STRAIN: T-100
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..530
; IDENTIFICATION METHOD: experimentally
US-08-696-834-1

Query Match 85.7%; Score 18; DB 2; Length 530;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 364 EAGVTS 369

RESULT 35
US-08-942-673-1
; Sequence 1, Application US/08942673
; Patent No. 5861292
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mineo
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshinori
```

```
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshinori
; APPLICANT: Yoshida, Masaru
; APPLICANT: Suzuki, Hiromi
; TITLE OF INVENTION: No. 5861292el L-sorbose Dehydrogenase and No. 5861292el
; TITLE OF INVENTION: L-sorbose Dehydrogenase Obtained from Gluconobacter
; TITLE OF INVENTION: oxydans T-100
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
; STREET: 1755 Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-DOS Editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/942,673
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513,841
; FILING DATE: 01-NOV-1995
; APPLICATION NUMBER: UK 9304700.9
; FILING DATE: 08-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 241851/1993
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-909-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Gluconobacter oxydans
; STRAIN: T-100
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..530
; IDENTIFICATION METHOD: experimentally
US-08-942-673-1

Query Match 85.7%; Score 18; DB 2; Length 530;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 364 EAGVTS 369

RESULT 36
US-09-118-317-1
; Sequence 1, Application US/09118317
; Patent No. 6197562
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mineo
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshinori
```


APPLICANT: Yoshida, Masaru
APPLICANT: Suzuki, Hiroki
TITLE OF INVENTION: No. 6197562el L-sorbose Dehydrogenase and No. 6197562el
TITLE OF INVENTION: L-sorbose Dehydrogenase Obtained from Gluconobacter
TITLE OF INVENTION: oxydants T-100
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Obion, Spivak, McClelland, Maier & Neustadt, P.C.
STREET: 1755 Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS Editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/118,317
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/513,841
FILING DATE: 01-NOV-1995
APPLICATION NUMBER: UK 9304700.9
FILING DATE: 08-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 241851/1993
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-909-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Gluconobacter oxydants
STRAIN: T-100
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1-530
IDENTIFICATION METHOD: experimentally
US-09-118-317-1

Query Match 85.7%; Score 18; DB 4; Length 530;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
Db 364 EAGVTS 369

RESULT 37
US-08-808-931-18
Sequence 18, Application US/08808931
Patent No. 5939602
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Heifetz, Peter

TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 5939602artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-931-18

Query Match 85.7%; Score 18; DB 2; Length 560;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
Db 67 EAGSGS 72

RESULT 38
US-08-808-323-18
Sequence 18, Application US/08808323
Patent No. 6018105
GENERAL INFORMATION:
APPLICANT: Johnson, Marie
APPLICANT: Volrath, Sandra
APPLICANT: Ward, Eric
TITLE OF INVENTION: Promoters from Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6018105artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,323
FILING DATE: 28-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-323-18

Query Match 85.7%; Score 18; DB 3; Length 560;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 67 EAGSGS 72

RESULT 39
US-09-050-603A-18
Sequence 18, Application US/09050603A
Patent No. 6023012
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6023012artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,603A
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-050-603A-18

Query Match 85.7%; Score 18; DB 3; Length 560;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 67 EAGSGS 72

RESULT 40
US-09-102-420B-18
Sequence 18, Application US/09102420B
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
TITLE OF INVENTION: OXIDASE ("PROTOX")
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6084155artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,420B
FILING DATE: 22-JUN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/059,164
FILING DATE: 13-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-102-420B-18

Query Match 85.7%; Score 18; DB 3; Length 560;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 67 EAGSGS 72

RESULT 41
US-09-497-698-18
Sequence 18, Application US/09497698
Patent No. 6308458
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
Ward, Eric
Heifetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
OXIDASE ("PROTOX")
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6308458artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/497,698
FILING DATE: 03-Feb-2000
CLASSIFICATION: <Unknown>
30-MAR-1998
11-MAR-1998

28-FEB-1997
28-FEB-1996
28-FEB-1996
21-JUN-1996
06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/102,420
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
STRANDEDNESS: NO. 6308458 Relevant
TOPOLOGY: NO. 6308458 Relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-497-698-18

Query Match 85.7%; Score 18; DB 4; Length 560;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 67 EAGSGS 72

RESULT 42
US-08-419-078-2
Sequence 2, Application US/08419078
Patent No. 5587306
GENERAL INFORMATION:
APPLICANT: HAWKINS, PHILLIP R.
APPLICANT: SEILHAMER, JEFFREY J.
TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3330 HILLVIEW AVENUE
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,078
FILING DATE:

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0030 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-855-0572
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; LIBRARY: No. 5587306e
; CLONE: 9118
; US-08-419-078-2

Query Match 85.7%; Score 18; DB 1; Length 566;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 338 EAGQS 343

RESULT 43
US-08-726-883-2
; Sequence 2, Application US/08726883
; Patent No. 5676946
; GENERAL INFORMATION:
; APPLICANT: HAWKINS, PHILLIP R.
; APPLICANT: SEILHAMER, JEFFREY J.
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILLVIEW AVENUE
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,883
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/419,078
; FILING DATE: 10-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0030 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-855-0572
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:

; LIBRARY: No. 5676946e
; CLONE: 9118
; US-08-726-883-2

Query Match 85.7%; Score 18; DB 1; Length 566;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 338 EAGQS 343

RESULT 44
US-08-696-944-2
; Sequence 2, Application US/08696944
; Patent No. 5981831
; GENERAL INFORMATION:
; APPLICANT: Sumant CHENGAPPA
; APPLICANT: Susan A. HELLYER
; APPLICANT: John S. REID
; APPLICANT: Jacqueline DE SILVA
; TITLE OF INVENTION: No. 5981831el Exo-(1-4)-Beta-D Galactanase
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,944
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00372
; FILING DATE: 23-FEB-1995
; APPLICATION DATA:
; APPLICATION NUMBER: GB 9403423.8
; FILING DATE: 23-FEB-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-696-944-2

Query Match 85.7%; Score 18; DB 2; Length 730;
Best Local Similarity 66.7%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
Db 603 EAGSNS 608

RESULT 45
US-08-731-716-2
; Sequence 2, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rocky, Pamela K.
; APPLICANT: Zhao, Genshi

APPLICANT: Rosteck, Paul R. Jr.
TITLE OF INVENTION: Penicillin Binding Protein From
TITLE OF INVENTION: Streptococcus Pneumoniae
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,716
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-731-716-2

Query Match 85.7%; Score 18; DB 1; Length 731;
Best Local Similarity 66.7%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 102 EAGALS 107

RESULT 46
US-09-651-656-19
Sequence 19, Application US/09651656
Patent No. 6340566
GENERAL INFORMATION:
APPLICANT: McCUTHEN-MALONEY, SANDRA
TITLE OF INVENTION: LAWRENCE LIVERMORE NATIONAL LABORATORY
TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
FILE REFERENCE: IL-10689
CURRENT APPLICATION NUMBER: US/09/651,656
CURRENT FILING DATE: 2000-08-29
PRIOR FILING DATE: 60/192,764
NUMBER OF SEQ ID NOS: 106
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 19
LENGTH: 823
TYPE: PRT
ORGANISM: Homo sapiens
US-09-651-656-19

Query Match 85.7%; Score 18; DB 4; Length 823;
Best Local Similarity 66.7%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 321 EAGSGS 326
RESULT 47
US-08-434-730-14
Sequence 14, Application US/08434730
Patent No. 5637463
GENERAL INFORMATION:
APPLICANT: Dalton, Stephen
APPLICANT: Kochan, Jarema P
APPLICANT: Osborne, Mark A
TITLE OF INVENTION: METHOD TO DETECT PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTIONS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,730
FILING DATE: 04-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Semionow, Raina
REGISTRATION NUMBER: 39022
REFERENCE/DOCKET NUMBER: 9069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)235-4391
TELEFAX: (201)235-2363
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 968 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-434-730-14

Query Match 85.7%; Score 18; DB 1; Length 968;
Best Local Similarity 66.7%; Pred. No. 3.5e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 711 EAGVTS 716

RESULT 48
US-08-560-005-2
Sequence 2, Application US/08560005
Patent No. 6001354
GENERAL INFORMATION:
APPLICANT: Pot, David A.
APPLICANT: Williams, Lewis T.
APPLICANT: Jefferson, Anne Bennett
APPLICANT: Majerus, Philip W.
TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
TITLE OF INVENTION: Acids Encoding Therefor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart tower, Suite 2000

;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94105
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/560,005
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dow, Karen B.
;; REGISTRATION NUMBER: 29,684
;; REFERENCE/DOCKET NUMBER: 2307K-0624000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-326-2400
;; TELEFAX: 415-326-2422
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 976 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-560-005-2

Query Match 85.7%; Score 18; DB 3; Length 976;
Best Local Similarity 66.7%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 496 EAGVTS 501

RESULT 49
US-09-195-868-14
; Sequence 14, Application US/09195868
; Patent No. 6090621
; GENERAL INFORMATION:
; APPLICANT: KAVANAUGH MD, MICHAEL
; APPLICANT: POT PH.D., DAVID
; APPLICANT: WILLIAMS MDPHD, LEWIS T.
; TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
; TITLE OF INVENTION: 5-PHOSPHATASES (SIPS)
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 HORTON STREET
; CITY: EMERYVILLE
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,868
; FILING DATE:
; CLASSIFICATION:
; APPLICATION NUMBER: US/09/195,868
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FIRESTONE, LEIGH H.
; REGISTRATION NUMBER: 36,831
; REFERENCE/DOCKET NUMBER: 1182.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2707

;; TELEFAX: 510-655-3542
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 976 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-195-868-14

Query Match 85.7%; Score 18; DB 3; Length 976;
Best Local Similarity 66.7%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 496 EAGVTS 501

RESULT 50
US-09-418-540-2
; Sequence 2, Application US/09418540
; Patent No. 6296848
; GENERAL INFORMATION:
; APPLICANT: Pot, David A.
; APPLICANT: Williams, Lewis T.
; APPLICANT: Jefferson, Anne Bennett
; APPLICANT: Majerus, Philip W.
; TITLE OF INVENTION: No. 6296848el Grb2 Associating Protein and Nucleic
; TITLE OF INVENTION: Acids Encoding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/418,540
; FILING DATE: 14-OCT-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/560,005
; FILING DATE: 17-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0624000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-418-540-2

Query Match 85.7%; Score 18; DB 4; Length 976;
Best Local Similarity 66.7%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 496 EAGVTS 501

Search completed: September 24, 2002, 11:23:07
Job time: 4456 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 10:16:21 ; Search time 54.84 Seconds
(without alignments)
10.513 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	19	90.5	37	2 A48620	adhesin - Staphylo
2	19	90.5	83	2 AF2563	hypothetical prote
3	19	90.5	122	2 C84320	hypothetical prote
4	19	90.5	143	2 B72627	hypothetical prote
5	19	90.5	145	2 F82189	hypothetical prote
6	19	90.5	164	2 T11215	hypothetical prote
7	19	90.5	200	2 T29807	hypothetical prote
8	19	90.5	201	2 T23855	hypothetical prote
9	19	90.5	204	2 A70844	probable moaE3 pro
10	19	90.5	212	2 AD1560	two-component resp
11	19	90.5	212	2 AF1202	two-component resp
12	19	90.5	220	2 AH0459	Sec-independent pr
13	19	90.5	245	2 E84169	hypothetical prote
14	19	90.5	246	2 AE1029	probable exported
15	19	90.5	257	2 T48058	RING-H2 zinc finger
16	19	90.5	260	2 B90026	hypothetical prote
17	19	90.5	263	2 G87721	protein ZC123.3 [1
18	19	90.5	266	2 JCI071	coat protein - soy
19	19	90.5	267	2 S18931	coat protein - soy
20	19	90.5	302	2 H82638	hypothetical prote
21	19	90.5	303	1 D64070	ATP phosphoribosyl
22	19	90.5	311	2 A56235	transcription acti
23	19	90.5	322	2 S38091	hypothetical prote
24	19	90.5	323	2 I49529	transcription fact
25	19	90.5	332	2 A38873	myristylated alani
26	19	90.5	352	2 S16547	neutral proteinase
27	19	90.5	352	2 G95872	probable adenylate
28	19	90.5	357	2 T01571	hypothetical prote
29	19	90.5	363	2 A83177	probable N-acetyl

30	19	90.5	363	2 T47240	amino acid transpo
31	19	90.5	365	2 E83600	hypothetical prote
32	19	90.5	367	1 I40226	3-isopropylmalate
33	19	90.5	387	2 T25452	hypothetical prote
34	19	90.5	389	1 A39429	cAMP response elem
35	19	90.5	390	2 J70334	acid proteinase (E
36	19	90.5	393	2 F72068	dihydrolipoamide s
37	19	90.5	393	2 C86556	dihydrolipoamide s
38	19	90.5	397	2 AD1559	glycine betaine AB
39	19	90.5	397	2 AF1201	probable processin
40	19	90.5	403	2 B72778	glycine betaine AB
41	19	90.5	443	2 A54813	cAMP receptor CAR4
42	19	90.5	444	2 S35783	glycoprotein gX -
43	19	90.5	458	2 T02804	hypothetical prote
44	19	90.5	462	2 D85438	hypothetical prote
45	19	90.5	469	2 C87629	major facilitator
46	19	90.5	474	2 D70853	hypothetical prote
47	19	90.5	482	2 S08384	modulation protein
48	19	90.5	482	2 S10133	probable transcrip
49	19	90.5	490	2 T41039	probable membrane
50	19	90.5	492	2 S54536	

ALIGNMENTS

RESULT 1

A48620
adhesin - Staphylococcus aureus (fragment)
C:Species: Staphylococcus aureus
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A48620
R:Patti, J.M.; Boles, J.O.; Hook, M.
Biochemistry 32, 11428-11435, 1993
A:Title: Identification and biochemical characterization of the ligand binding domain
A:Reference number: A48620; MUID:94032261
A:Contents: FGA 574
A:Accession: A48620
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-37 <PAT>
A:Note: sequence extracted from NCBI backbone (NCBIP:138726)

Query Match 90.5%; Score 19; DB 2; Length 37;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 16 EAGTSS 21

RESULT 2

AF2563
hypothetical protein asl8505 [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC7
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120.
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AF2563
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2563
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <KUR>
A:Cross-references: GB:AP003604; PIDN:BA877424.1; PID:gl7134868; GSPDB:GN00183
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asl8505

A:Genome: plasmid

Query Match 90.5%; Score 19; DB 2; Length 83;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 78 EAGASS 83

RESULT 3

C84320

hypothetical protein Vng1678h [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84320

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: C84320

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <STO>

A:Cross-references: GB:AE004437; NID:g10581147; PIDN:AAG19927.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG1678H

Query Match 90.5%; Score 19; DB 2; Length 122;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 57 EAGASS 62

RESULT 4

B72627

hypothetical protein APE1474 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: B72627

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339

A:Accession: B72627

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-143 <RAW>

A:Cross-references: DDBJ:AF000061; NID:g5104821; PIDN:BAA80472.1; PID:d1044258; PID:g510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1474

Query Match 90.5%; Score 19; DB 2; Length 143;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 85 EAGASS 90

RESULT 5

F82189

hypothetical protein VC1536 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: F82189

R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: F82189

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-145 <HEI>

A:Cross-references: GB:AE004231; GB:AE003852; NID:g9656027; PIDN:AAF94690.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1536

A:Map position: 1

Query Match 90.5%; Score 19; DB 2; Length 145;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 114 EAGSTS 119

RESULT 6

T11215

hypothetical protein 5 - Streptomyces glaucescens

C:Species: Streptomyces glaucescens

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T11215

R:Summers, R.G.; Ali, A.; Shen, B.; Wessel, W.A.; Hutchinson, C.R.

Biochemistry 34, 9389-9402, 1995

A:Title: Malonyl-coenzyme A:acyl carrier protein acyltransferase of Streptomyces glau

A:Reference number: Z17254; MUID:95352622

A:Accession: T11215

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-164 <SUM>

A:Cross-references: EMBL:L43074; NID:g870805; PID:g870810

Query Match 90.5%; Score 19; DB 2; Length 164;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 55 EAGTAS 60

RESULT 7

T29807

hypothetical protein C25A8.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29807

R:Latreille, P.; Stellyes, L.

submitted to the EMBL Data Library, June 1996

A:Description: The sequence of C. elegans cosmid C25A8.

A:Reference number: Z20689

A:Accession: T29807

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-200 <LAT>

A:Cross-references: EMBL:U61958; PIDN:AAB03180.1; GSPDB:GN000022; CESP:C25A8.2

A;Experimental source: strain Bristol N2; clone C25A8
C;Genetics:
A;Gene: CESP:C25A8.2
A;Map position: 4
A;Introns: 173/3

Query Match 90.5%; Score 19; DB 2; Length 200;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 109 EAGNAS 114

RESULT 8
T23855

hypothetical protein R02D5.7 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T23855

R;Matthews, L.

submitted to the EMBL Data Library; August 1996

A;Reference number: Z19808

A;Accession: T23855

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-201 <WIL>

A;Cross-references: EMBL:Z78015; PIDN:CAB01436.1; GSPDB:GN00023; CESP:R02D5.7

A;Experimental source: clone R02D5

C;Genetics:

A;Gene: CESP:R02D5.7

A;Map position: 5

A;Introns: 174/3

Query Match 90.5%; Score 19; DB 2; Length 201;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 110 EAGAAS 115

RESULT 9

A70844

probable moaE3 protein - *Mycobacterium tuberculosis* (strain H37RV)

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C;Accession: A70844

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A;Reference number: A70500; MUID:98295987

A;Accession: A70844

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-204 <COL>

A;Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAAL7094.1; PID:e125113

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: moaE3

Query Match 90.5%; Score 19; DB 2; Length 204;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 153 EAGTAS 158

RESULT 10

AD1560

two-component response regulator, in particular B. subtilis YvqC protein homolog lml1
C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C;Accession: AD1560

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AD1560

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-212 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC96252.1; PID:g16413480; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lml021

C;Superfamily: regulatory protein comA; response regulator homology

Query Match 90.5%; Score 19; DB 2; Length 212;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 95 EAGASS 100

RESULT 11

AF1202

two-component response regulator, in particular B. subtilis YvqC protein homolog lml1
C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C;Accession: AF1202

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AF1202

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-212 <GLA>

A;Cross-references: GB:NC_003210; PIDN:CAC99100.1; PID:g16410424; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lml01022

C;Superfamily: regulatory protein comA; response regulator homology

Query Match 90.5%; Score 19; DB 2; Length 212;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 95 EAGASS 100

RESULT 12
AH0459

Sec-independent protein translocase protein TatB [imported] - Yersinia pestis (strain CO
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0459
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AH0459

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-220 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC93244.1; PID:g15981690; GSPDB:GN00175

C;Genetics:

A;Gene: tatB

Query Match

Best Local Similarity 90.5%; Score 19; DB 2; Length 220;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6

||| |

Db 143 EAGTAS 148

RESULT 13

E84169

hypothetical protein p1mT1 [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-feb-2001 #text_change 02-Feb-2001

C;Accession: E84169

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483

A;Accession: E84169

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-245 <STO>

A;Cross-references: GB:AE004437; NID:g10579741; PIDN:AAG18721.1; GSPDB:GN00138

C;Genetics:

A;Gene: p1mT1

Query Match

Best Local Similarity 90.5%; Score 19; DB 2; Length 245;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6

||| |

Db 151 EAGAS 156

RESULT 14

AE1029

Probable exported protein STX458 [imported] - Salmonella enterica subsp. enterica serov

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C;Accession: AE1029

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A;Reference number: AB0502; PMID:11677608

A;Accession: AE1029

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-246 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD09334.1; PID:g16505334; GSPDB:GN00176

C;Genetics:

A;Gene: STY4558

Query Match

Best Local Similarity 90.5%; Score 19; DB 2; Length 246;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6

||| |

Db 93 EAGSAS 98

RESULT 15

T48058

RING-H2 zinc finger protein ATL5 - Arabidopsis thaliana

N;Alternate names: protein F26K9.120

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 08-Dec-2000

C;Accession: T48058

R;Bloeker, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salano

submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24465

A;Accession: T48058

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-257 <BLO>

A;Cross-references: EMBL:AL162651

A;Experimental source: cultivar Columbia; BAC clone F26K9

C;Genetics:

A;Map position: 3

A;Note: F26K9.120

C;Superfamily: Arabidopsis hypothetical protein F19I3.22; RING finger homology

F;109-160/Domain: RING finger homology <RRN>

Query Match

Best Local Similarity 90.5%; Score 19; DB 2; Length 257;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6

||| |

Db 184 EAGSS 189

RESULT 16

B90026

hypothetical protein modA [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C;Accession: B90026

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A;Reference number: A99759; MUID:21311952; PMID:11418146

A;Accession: B90026

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-260 <KUR>

A;Cross-references: GB:BA000018; PID:g13702079; PIDN:BA043371.1; GSPDB:GN00149

A;Experimental source: strain N315

C;Genetics:

A;Gene: modA

C;Superfamily: molybdate-binding periplasmic protein

Query Match 90.5%; Score 19; DB 2; Length 260;
 Best Local Similarity 66.7%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 227 EAGATS 232

RESULT 17
 G87721
 protein ZC123.3 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: G87721
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 35, 1999
 A:Accession: G87721
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-263 <STO>
 A:Cross-references: GB:chr_I; PIDN:AAB97603.1; PID:g2804499; GSPDB:GN00019; CESP:ZC123.3
 A:Note: contains similarity to C2H2-type zinc fingers
 C:Genetics:
 A:Gene: ZC123.3
 A:Map position: 1

Query Match 90.5%; Score 19; DB 2; Length 263;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 31 EAGASS 36

RESULT 18
 JC1071
 coat protein - soybean mosaic virus
 C:Species: soybean mosaic virus, SbMV
 C:Date: 02-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 17-Nov-2000
 C:Accession: JC1071
 R:Chu, R.Y.; Leng, X.H.; Bao, Y.M.; Pan, N.S.; Pu, Z.Q.
 Acta Bot. Sin. 34, 523-528, 1992
 A:Title: Amplification of soybean mosaic virus coat protein gene by polymerase chain reaction
 A:Reference number: JC1071
 A:Accession: JC1071
 A:Molecule type: DNA
 A:Residues: 1-266 <CHU>
 C:Superfamily: tobacco etch virus genome polyprotein

Query Match 90.5%; Score 19; DB 2; Length 266;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 25 EAGTSS 30

RESULT 19
 S18931
 coat protein - soybean mosaic virus (fragment)
 C:Species: soybean mosaic virus, SbMV
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001

Query Match 90.5%; Score 19; DB 2; Length 267;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 26 EAGTSS 31

RESULT 20
 H82638
 hypothetical protein XF1783 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: H82638
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: H82638
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <SIM>
 A:Cross-references: GB:AE004000; GB:AE003849; NID:g9106850; PIDN:AAF84591.1; GSPDB:GN
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
 as-Neto, E.; Docena, C.; El-Dorriy, H.; Pacincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF1783

Query Match 90.5%; Score 19; DB 2; Length 302;
 Best Local Similarity 66.7%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 106 EAGTAS 111

RESULT 21
 D64070
 ATP phosphoribosyltransferase (EC 2.4.2.17) - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Mar-2000
 C:Accession: D64070
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

C:Accession: S18931
 R:Chu, R.
 submitted to the EMBL Data Library, January 1992
 A:Description: cDNA sequence of the gene encoding coat protein of SMV.
 A:Reference number: S18931
 A:Accession: S18931
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <CHU>
 A:Cross-references: EMBL:X63771; NID:g61983; PIDN:CAA45307.1; PID:g61984
 C:Superfamily: tobacco etch virus genome polyprotein

Query Match 90.5%; Score 19; DB 2; Length 267;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 26 EAGTSS 31

RESULT 20
 H82638
 hypothetical protein XF1783 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: H82638
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: H82638
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <SIM>
 A:Cross-references: GB:AE004000; GB:AE003849; NID:g9106850; PIDN:AAF84591.1; GSPDB:GN
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
 as-Neto, E.; Docena, C.; El-Dorriy, H.; Pacincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF1783

Query Match 90.5%; Score 19; DB 2; Length 302;
 Best Local Similarity 66.7%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 106 EAGTAS 111

RESULT 21
 D64070
 ATP phosphoribosyltransferase (EC 2.4.2.17) - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Mar-2000
 C:Accession: D64070
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: D64070
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-303 <TIGR>
A:Cross-references: GB:U32729; GB:I42023; NID:gl573439; PIDN:AAC22127.1; PID:gl573446; J
A:Note: named as homolog to a protein from Escherichia coli
C:Superfamily: ANP phosphoribosyltransferase; ATP phosphoribosyltransferase homology
C:Keywords: glycosyltransferase; histidine biosynthesis; pentosyltransferase

Query Match 90.5%; Score 19; DB 1; Length 303;
Best Local Similarity 66.7%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 287 EAGASS 292

RESULT 22
A56235
transcription activator MafB - chicken
C:Species: Gallus gallus (chicken)
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
C:Accession: A56235
R:Kataoka, K.; Fujiwara, K.T.; Noda, M.; Nishizawa, M.
Mol. Cell. Biol. 14, 7581-7591, 1994
A:Title: MafB, a new Maf family transcription activator that can associate with Maf and
A:Reference number: A56235; MUID:95021288
A:Accession: A56235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <KAT>
A:Cross-references: GB:D28600; NID:g516723; PIDN:BAA05938.1; PID:g516724
C:Genetics:

Query Match 90.5%; Score 19; DB 2; Length 311;
Best Local Similarity 66.7%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 296 EAGSTS 301

RESULT 23
S38091
hypothetical protein YKR022c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Oct-1999
C:Accession: S38091
R:Duesterhoeft, A.; Moestl, D.; Poehlmann, R.; Philippssen, P.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37811
A:Accession: S38091
A:Molecule type: DNA
A:Residues: 1-322 <DUE>
A:Cross-references: EMBL:Z28247; NID:g486444; PID:g486445; GSPDB:GN00011; MIPS:YKR022c
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YKR022c
A:Map position: 11R

Query Match 90.5%; Score 19; DB 2; Length 322;
Best Local Similarity 66.7%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 119 EAGSSS 124

RESULT 24
I49529
transcription factor-kr - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 16-Jul-1999
C:Accession: I49529
R:Cordes, S.P.; Barsh, G.S.
Cell 79, 1025-1034, 1994
A:Title: The mouse segmentation gene kr encodes a novel basic domain-leucine zipper t
A:Reference number: A55200; MUID:95094266
A:Accession: I49529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-323 <RES>
A:Cross-references: GB:L36435; NID:g625043; PIDN:AAA65689.1; PID:g625044
C:Superfamily: maf transforming protein; maf homology
C:Keywords: leucine zipper; transcription factor
F:212-301/Domain: maf homology <MAF>

Query Match 90.5%; Score 19; DB 2; Length 323;
Best Local Similarity 66.7%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 308 EAGSTS 313

RESULT 25
A38873
myristylated alanine-rich protein kinase C substrate - human
N:Alternate names: acidic calmodulin-binding 80K protein; MARCKS
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C:Accession: A38873; A42977; A40758; S29269
R:Shimizu, N.
submitted to DDBJ, September 1991
A:Reference number: A38873
A:Accession: A38873
A:Molecule type: mRNA
A:Residues: 1-332 <SHI>
A:Cross-references: GB:D10522; GB:D90498; NID:g219893; PIDN:BAA01392.1; PID:g219894
R:Sakai, K.; Hirai, M.; Kudo, J.; Minoshima, S.; Shimizu, N.
Genomics 14, 175-178, 1992
A:Title: Molecular cloning and chromosomal mapping of a cDNA encoding human 80K-L pro
A:Reference number: A42977; MUID:93052291
A:Accession: A42977
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-13, G', 15-332 <SAK>
A:Cross-references: GB:D90498
A:Experimental source: squamous carcinoma cells A431
A:Note: sequence extracted from NCBI backbone (NCBIP:118653)
R:Harlan, D.M.; Graff, J.M.; Stumpo, D.J.; Eddy Jr., R.L.; Shows, T.B.; Boyle, J.M.;
J. Biol. Chem. 266, 14399-14405, 1991
A:Title: The human myristoylated alanine-rich C kinase substrate (MARCKS) gene (MACS)
A:Reference number: A40758; MUID:91317795
A:Accession: A40758
A:Molecule type: mRNA
A:Residues: 1-83, 'A', 85-118, 'P', 120-233, 'W', 235-286, 'LYC', 290, 'RRGSGPRGGARRSLNQ', 30
A:Cross-references: GB:M68956

A:Note: the authors translated the codon GGC for residue 53 as Arg
 R;Herget, T.; Brooks, S.F.; Broad, S.; Rozenfurt, E.
 Eur. J. Biochem. 209, 7-14, 1992
 A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein
 or equivalent genes in different species.
 A:Reference number: S29267; MUID:93011168
 A:Accession: S29269
 A:Molecule type: mRNA
 A:Residues: 189-223, R, 225-234, E, 236-322 <HER>
 C:Comment: This protein is a major cellular substrate for protein kinase C and plays a
 C:Comment: It binds to calmodulin in one to one molar ratio in the presence of calcium
 C:Genetics:
 A:Gene: GDB:WACS
 A:Cross-references: GDB:118835; OMIM:177061
 A:Map position: 6q22.2-6q22.2
 C:Superfamily: neurofilament triplet H protein
 C:Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristyl
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F;159,163,167,170/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status

Query Match 90.5%; Score 19; DB 2; Length 332;
 Best Local Similarity 66.7%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
 ||| |
 Db 206 EAGAAS 211

RESULT 26

S16547
 Neutral proteinase II - Aspergillus oryzae
 C:Species: Aspergillus oryzae
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
 C:Accession: S16547; S47562
 R;Tatsumi, H.; Murakami, S.; Tsuji, R.F.; Ishida, Y.; Murakami, K.; Masaki, A.; Kawabe,
 Mol. Gen. Genet. 228, 97-103, 1991
 A:Title: Cloning and expression in yeast of a cDNA clone encoding Aspergillus oryzae neu
 A:Reference number: S16547; MUID:91360097
 A:Accession: S16547
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-352 <TAT>
 A:Cross-references: GB:S53810; NID:9234832; PIDN:AB19701.1; PID:g234833
 R;Tatsumi, H.; Ikegaya, K.; Murakami, S.; Kawabe, H.; Nakano, E.; Motai, H.
 Biochim. Biophys. Acta 1208, 179-185, 1994
 A:Title: Elucidation of the thermal stability of the neutral proteinase II from Aspergil
 A:Reference number: S47562; MUID:94368822
 A:Accession: S47562
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 176-352 <TA2>
 C:Superfamily: Penicillium citrinum penicillolysin

Query Match 90.5%; Score 19; DB 2; Length 352;
 Best Local Similarity 66.7%; Pred. No. 7.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
 ||| |
 Db 240 EAGSTS 245

RESULT 27

G95872
 Probable adenylate cyclase protein [imported] - Sinorhizobium meliloti (strain 1021) mag
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: G95872
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9389-9394, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: G95872
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-352 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48647.1; PID:g15140119; GSPDB:GN00167
 A:Experimental source: Strain 1021, megaplasmid pSymB
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
 heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMD20257
 A:Genome: plasmid

Query Match 90.5%; Score 19; DB 2; Length 352;
 Best Local Similarity 66.7%; Pred. No. 7.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
 ||| |
 Db 13 EAGTSS 18

RESULT 28

T01571
 Hypothetical protein A_TM018A10.10 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Nov-1999
 C:Accession: T01571
 R;Dempsey, S.; Harper, M.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of A. thaliana TM018A10.
 A:Reference number: 214348
 A:Accession: T01571
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-357 <DEM>
 A:Cross-references: EMBL:AF013294; NID:g2252848; PID:g2252871
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Map position: 4
 A:Introns: 47/3; 201/2; 243/1; 259/2
 A:Note: A_TM018A10.10
 C:Superfamily: Arabidopsis thaliana hypothetical protein A_TM018A10.10

Query Match 90.5%; Score 19; DB 2; Length 357;
 Best Local Similarity 66.7%; Pred. No. 7.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
 ||| |
 Db 91 EAGSSS 96

RESULT 29

A83177
 Probable N-acetylglucosamine-6-phosphate deacetylase PA3758 [imported] - Pseudomonas
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: A83177
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: AB2950; MUID:20437337

A:Accession: AB3177
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-363 <STO>
 A:Cross-references: GB:AB004794; GB:AB004091; NID:g9949917; PIDN:AA007145.1; GSPDB:GN001

A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3758

Query Match 90.5%; Score 19; DB 2; Length 363;
 Best Local Similarity 66.7%; Pred. No. 7.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 |||||
 Db 196 EAGASS 201

RESULT 30
 T47240

amino acid transport protein arg-1, mitochondrial [imported] - *Neurospora crassa*
 C:Species: *Neurospora crassa*

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
 C:Accession: T47240

R:Liu, Q.; Dunlap, J.C.

Genetics 143, 1183-1174, 1996

A:Title: Isolation and analysis of the arg-13 gene of *Neurospora crassa*.

A:Reference number: Z24416; MUID:96400914

A:Accession: T47240

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-363 <LIU>

A:Cross-references: EMBL:L36378; NID:g773383; PIDN:AA037500.1; PID:g773384

A:Experimental source: strain bda; isolate 30-1

C:Genetics:

A:Gene: arg-13

A:Map position: V

A:Introns: 50/3

C:Keywords: amino acid transport; mitochondrion

Query Match 90.5%; Score 19; DB 2; Length 363;
 Best Local Similarity 66.7%; Pred. No. 7.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 |||||
 Db 14 EAGAA 19

RESULT 31
 E83800

hypothetical protein BH1205 [imported] - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: E83800

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: AB3650; MUID:20512582; PMID:11058132

A:Accession: E83800

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-365 <STO>

A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA04924.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1205

Query Match 90.5%; Score 19; DB 2; Length 387;
 Best Local Similarity 66.7%; Pred. No. 8.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 |||||
 Db 276 EAGSSS 281

Query Match 90.5%; Score 19; DB 2; Length 365;
 Best Local Similarity 66.7%; Pred. No. 7.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 |||||
 Db 247 EAGTAS 252

RESULT 32
 I40226

3-isopropylmalate dehydrogenase (EC 1.1.1.1.85) - *Bacillus megaterium*
 C:Species: *Bacillus megaterium*

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: I40226; S38506

R:Meinhardt, F.; Wittchen, K.D.; Buakamp, M.

Appl. Microbiol. Biotechnol. 41, 344-351, 1994

A:Title: Cloning and sequencing of the *lenc* and *npM* genes and a putative *spoIV* gene f

A:Reference number: I40226; MUID:94288995

A:Accession: I40226

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-367 <RES>

A:Cross-references: EMBL:X65184; NID:g414096; PIDN:CAA46295.1; PID:g414097

A:Experimental source: DSM 319

C:Genetics:

A:Gene: *lenc*

C:Function:

A:Pathway: leucine biosynthesis

C:Superfamily: 3-isopropylmalate dehydrogenase

C:Keywords: leucine biosynthesis; NAD; oxidoreductase

Query Match 90.5%; Score 19; DB 1; Length 367;
 Best Local Similarity 66.7%; Pred. No. 7.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 |||||
 Db 48 EAGSSS 53

RESULT 33
 T25452

hypothetical protein B0412.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T25452

R:Bentley, D.

submitted to the EMBL Data Library, December 1996

A:Description: The sequence of *C. elegans* cosmid B0412.

A:Reference number: Z20037

A:Accession: T25452

A:Status: preliminary; translated from GB/EMBL/DBDJ.

A:Molecule type: DNA

A:Residues: 1-387 <BEN>

A:Cross-references: EMBL:U80953; PIDN:AA052553.1; GSPDB:GN00021; CESP:B0412.1

A:Experimental source: strain Bristol N2; clone B0412

C:Genetics:

A:Gene: CESP:B0412.1

A:Map position: 3

A:Introns: 110/3; 146/2; 175/1; 213/2; 253/3; 318/1

RESULT 34

A39429
CAMP response element-binding protein ATF2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39429
R:Kageyama, R.; Sasai, Y.; Nakanishi, S.
J. Biol. Chem. 266, 15525-15531, 1991
A:Title: Molecular characterization of transcription factors that bind to the CAMP response element
A:Reference number: A39429; MUID:91332085
A:Accession: A39429
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-389 <RAG>
A:Cross-references: GB:M65148; NID:g206569; PIDN:AAA42013.1; PID:g206570
C:Superfamily: CAMP response element-binding protein 1; fos/jun DNA-binding domain homologues
C:Keywords: DNA binding; nucleus; transcription regulation
F:231-271/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match

Best Local Similarity 90.5%; Score 19; DB 1; Length 389;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
||| |
DB 354 EAGATS 359

RESULT 35

JT0334
acid proteinase (EC 3.4.23.-) PEPI precursor - yeast (Saccharomycopsis fibuligera)
C:Species: Saccharomycopsis fibuligera
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Sep-1999
C:Accession: JT0334
R:Hirata, D.; Fukui, S.; Yamashita, I.
Agric. Biol. Chem. 52, 2647-2649, 1988
A:Title: Nucleotide sequence of the secreted acid protease gene PEPI in the yeast saccharomyces cerevisiae
A:Reference number: JT0334
A:Accession: JT0334
A:Molecule type: DNA
A:Residues: 1-390 <HIR>
C:Genetics:
A:Gene: PEPI
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-390/Product: acid proteinase #status predicted <MAT>
F:93,282/Active site: Asp #status predicted

Query Match

Best Local Similarity 90.5%; Score 19; DB 1; Length 389;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
||| |
DB 354 EAGATS 359

RESULT 36

F72068
dihydrolipoamide succinyltransferase - Chlamydomonas reinhardtii (strain CWL029)
C:Species: Chlamydomonas reinhardtii (strain CWL029)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: F72068
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydomonas reinhardtii and C. trachomatis.
A:Reference number: A72068; MUID:99206606
A:Accession: F72068
A:Status: preliminary
A:Molecule type: DNA

Query Match

Best Local Similarity 90.5%; Score 19; DB 2; Length 390;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
||| |
DB 266 EAGSSS 271

RESULT 37

C86556
dihydrolipoamide succinyltransferase [imported] - Chlamydomonas reinhardtii (strain J1)
C:Species: Chlamydomonas reinhardtii (strain J1)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: C86556
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii J138.
A:Reference number: A86491; MUID:20330349
A:Accession: C86556
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:BA000008; NID:g8978898; PIDN:BAA98733.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: sucB.2
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match

Best Local Similarity 90.5%; Score 19; DB 2; Length 393;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
||| |
DB 101 EAGSSS 106

RESULT 38

AD1559
glycine betaine ABC transporter (ATP-binding protein) homolog gbuA [imported] - Listeria innocua
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1559
R:Glaser, P.; Franjeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96244.1; PID:g16413472; GSPDB:GN00178
A:Experimental source: strain Clip1262
C:Genetics:
A:Gene: gbuA
C:Superfamily: glycine betaine/proline transport protein prov; ATP-binding cassette h

Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 301 EAGTSS 306

RESULT 39
AF1201
glycine betaine ABC transporter (ATP-binding protein) homolog gbuA [imported] - Listeria
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1201
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1201
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <GIA>
A:Cross-references: GB:NC_003210; PIDN:CAC99092.1; PID:g16410416; GSPDB:GN00177
C:Genetics:
A:Gene: gbuA
C:Superfamily: glycine betaine/proline transport protein prov; ATP-binding cassette hom

Query Match 90.5%; Score 19; DB 2; Length 397;
Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 301 EAGTSS 306

RESULT 40
B72778
probable processing proteinase APE0212 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: B72778
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339
A:Accession: B72778
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-403 <KAW>
A:Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA79124.1; PID:g5103603
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0212
C:Superfamily: mitochondrial processing peptidase alpha chain

Query Match 90.5%; Score 19; DB 2; Length 403;
Best Local Similarity 66.7%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 266 EAGATS 271

RESULT 41
A54813

CAMP receptor CAR4 - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 07-May-1999
C:Accession: A54813
R:Louis, J.M.; Ginsburg, G.T.; Kimmel, A.R.
Genes Dev. 8, 2086-2096, 1994

A:Title: The CAMP receptor CAR4 regulates axial patterning and cellular differentiati
A:Reference number: A54813; MUID:95047357
A:Accession: A54813
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: DNA
A:Residues: 1-443 <LOU>
C:Genetics:
A:Gene: CAR4
C:Keywords: CAMP binding

Query Match 90.5%; Score 19; DB 2; Length 443;
Best Local Similarity 66.7%; Pred. No. 9.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 72 EAGSTS 77

RESULT 42

S35783
glycoprotein gX - bovine herpesvirus 1
C:Species: bovine herpesvirus 1
C:Date: 09-Jun-1994 #sequence_revision 12-May-1995 #text_change 08-Oct-1999
C:Accession: S35783
R:Audonnet, J.
submitted to the EMBL Data Library, June 1993
A:Reference number: S35782
A:Accession: S35783
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <AUD>
A:Cross-references: EMBL:223068; NID:g312185; PIDN:CAA80603.1; PID:g312187
C:Superfamily: pseudorabies virus glycoprotein gX
C:Keywords: glycoprotein

Query Match 90.5%; Score 19; DB 2; Length 444;
Best Local Similarity 66.7%; Pred. No. 9.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 267 EAGSAS 272

RESULT 43

T02804
hypothetical protein L2602.2 [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C:Accession: H81456; T02804
R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protei
A:Reference number: A81455; MUID:99178987
A:Accession: H81456
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <PYL>
A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24627.1; PID:g2995580; GSPDB:G
A:Experimental source: strain MHOM/IL/81/Friedlin
C:Genetics:
A:Gene: L2602.2

A:Map position: 1
C:Superfamily: Leishmania major hypothetical protein L2602.2

Query Match 90.5%; Score 19; DB 2; Length 458;
Best Local Similarity 66.7%; Pred. No. 9.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 10 EAGTAS 15

RESULT 44

D85438
hypothetical protein AT4g37110 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: D85438
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488

A:Accession: D85438
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <STO>
A:Cross-references: GB:NC_001268; NID:g7270660; PIDN:CAB80377.1; GSPDB:GN00140
A:Gene: AT4g37110
A:Map position: 4

Query Match 90.5%; Score 19; DB 2; Length 462;
Best Local Similarity 66.7%; Pred. No. 9.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 108 EAGAAS 113

RESULT 45

C87629
major facilitator family transporter CC3069 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: C87629
R:Nierman, W.C.; Feldblyum, T.V.; Faulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87629
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <STO>
A:Cross-references: GB:AE005673; NID:g13424719; PIDN:AAK25031.1; GSPDB:GN00148
A:Gene: CC3069
C:Superfamily: lincomycin-resistance protein lmrB

Query Match 90.5%; Score 19; DB 2; Length 469;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 394 EAGAAS 399

RESULT 46

D70853
hypothetical protein RV3088 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70853
R:Collier, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: D70853
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-474 <COL>
A:Cross-references: GB:AL021309; GB:AL123456; NID:g3261510; PIDN:CAA16146.1; PID:e124
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3088

Query Match 90.5%; Score 19; DB 2; Length 474;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 460 EAGTTS 465

RESULT 47

S08384
nodulation protein nodT - Rhizobium leguminosarum plasmid pIJ1089
C:Species: Rhizobium leguminosarum
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 18-Sep-1998
C:Accession: S08384
R:Economou, A.; Hamilton, W.D.O.; Johnston, A.W.B.; Downie, J.A.
EMBO J. 9, 349-354, 1990
A:Title: The Rhizobium nodulation gene nodO encodes a Ca²⁺-binding protein that is ex
A:Reference number: S08384; MUID:90151607
A:Accession: S08384
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-482 <ECO>
A:Cross-references: EMBL:X17285
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1989
C:Genetics:
A:Genome: plasmid pIJ1089
C:Superfamily: nodulation protein nodT

Query Match 90.5%; Score 19; DB 2; Length 482;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 201 EAGAAS 206

RESULT 48

S10133
nodulation protein nodT - Rhizobium leguminosarum bv. viciae
C:Species: Rhizobium leguminosarum bv. viciae
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
C:Accession: S10133
R:Surin, B.P.; Watson, J.M.; Hamilton, W.D.O.; Economou, A.; Downie, J.A.
Mol. Microbiol. 4, 245-252, 1990
A:Title: Molecular characterization of the nodulation gene, nodT, from two biovars of
A:Reference number: S08616; MUID:90251164
A:Accession: S10133

A:Molecule type: DNA
A:Residues: 1-482 <SUR>
A:Cross-references: EMBL:X17285; NID:g46251; PIDN:CAA351177.1; PID:g581512
C:Genetics:
A:Gene: nodT
A:Start codon: TTG
C:Superfamily: nodulation protein nodT

Query Match 90.5%; Score 19; DB 2; Length 482;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 201 EGAAS 206

RESULT 49

T41039
Probable transcription initiation factor Iif, alpha subunit - fission yeast (Schizosacch
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
C:Accession: T41039
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL data Library, December 1998
A:Reference number: Z21966
A:Accession: T41039
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-490 <LYN>
A:Cross-references: EMBL:AL034491; PIDN:CAA22493.1; GSPDB:GN00068; SPDB:SPCC1620.09c
A:Experimental source: strain 972h-; cosmid c1620
C:Genetics:
A:Gene: SPDB:SPCC1620.09c
A:Map position: 3
A:Introns: 180/3; 212/3
C:Keywords: transcription initiation

Query Match 90.5%; Score 19; DB 2; Length 490;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 234 EAGSAS 239

RESULT 50

S54536
Probable membrane protein YDR240c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YD8419.07c
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 04-Mar-2000
C:Accession: S54536
R:Oliver, K.; Harris, D.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54530
A:Accession: S54536
A:Molecule type: DNA
A:Residues: 1-492 <OLI>
A:Cross-references: EMBL:Z49701; NID:g817819; PID:g817826; GSPDB:GN00004; MIPS:YDR240c
A:Experimental source: strain AB972
C:Genetics:
A:Gene: MIPS:YDR240c
A:Map position: 4R
C:Superfamily: Saccharomyces cerevisiae probable membrane protein YDR240c
C:Keywords: transmembrane protein
F:119-135/Domain: transmembrane #status predicted <TMM>

Query Match 90.5%; Score 19; DB 2; Length 492;

Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 325 EAGATS 330

Search completed: September 24, 2002, 11:24:17
Job time: 4076 sec

THIS PAGE BLANK (USPTO)

OM of: BASK-853-CLAIM4 to: GenEmbl.* out_format : pfs

Date: Sep 24, 2002 12:09 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+bin.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/BASKAR853079/runat_17092002_103407_1819/app_query.fasta_1.98
-DB=GenEmbl -QMT=fastap -SUFFIX=mod.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS=human40.cdi -LIST=50 -DALIGN=200 -THR_SCORE=pet
-THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=BASKAR853079 @CGNL1.1.6129 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: BASK-853-CLAIM4

Query length: 6

Database: GenEmbl.*

Database sequences: 1797656

Database length: 1873333701

Search time (sec): 1991.950000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_pat:AU9389	+	19.00	108.04	1.2e+03	28 ! A19389 oligonucleotide. 5/1994
gb_pat:AR123943	-	19.00	107.25	1.3e+03	32 ! AR123943 Sequence 28 from patent
gb_pat:AR009555	-	19.00	106.90	1.4e+03	34 ! AR009555 Sequence 20 from patent
gb_pat:AR009593	+	19.00	106.90	1.4e+03	34 ! AR009593 Sequence 58 from patent
gb_pat:AU9413	+	19.00	106.56	1.4e+03	36 ! A19413 oligonucleotide. 6/1994
gb_pat:AU9414	+	19.00	106.56	1.4e+03	36 ! A19414 oligonucleotide. 6/1994
gb_pat:AU9438	+	19.00	106.56	1.4e+03	36 ! A19438 Sequence 41 from patent
gb_pat:AU9439	+	19.00	106.56	1.4e+03	36 ! A19439 Sequence 42 from patent
gb_pat:HSU30392	+	19.00	105.64	1.6e+03	42 ! U30392 Human isolate M103 T-cell
gb_pat:AX162903	+	19.00	104.50	1.8e+03	51 ! AX162903 Sequence 6231 from Pat
gb_pat:AX162904	+	19.00	104.50	1.8e+03	51 ! AX162904 Sequence 6232 from Pat
gb_pat:AX162905	+	19.00	104.50	1.8e+03	51 ! AX162905 Sequence 6233 from Pat
gb_pat:AX162906	+	19.00	104.50	1.8e+03	51 ! AX162906 Sequence 6234 from Pat
gb_pat:AF011601	+	19.00	104.50	1.8e+03	51 ! AF011601 Homo sapiens T cell rec
gb_pat:AR053957	+	19.00	101.54	2.7e+03	84 ! AR053957 Sequence 94 from patent
gb_in:AF185208	+	19.00	98.83	3.8e+03	133 ! AF185208 Pista australis U2 sma
gb_vi:HV8417	+	19.00	98.48	4.0e+03	141 ! AJ008417 Human immunodeficiency
gb_vi:HV8341	+	19.00	98.23	4.1e+03	147 ! AJ008341 Human immunodeficiency
gb_ov:CLDABC	-	19.00	97.88	4.3e+03	156 ! X9594 C.livia gene encoding al
gb_ov:FEERABC	-	19.00	97.88	4.3e+03	156 ! X9596 E.elegans mRNA for alpha
gb_ro:RN078130	-	19.00	97.84	4.3e+03	157 ! U78130 Rattus norvegicus zinc f
gb_ro:MWJ004556	-	19.00	97.66	4.4e+03	162 ! AJ004556 Maxomys whiteheadi M
gb_sts:DM62C3T	-	19.00	97.51	4.5e+03	166 ! Z32221 D. melanogaster STS dete
gb_sts:HUMSWS229	-	19.00	97.34	4.6e+03	171 ! L10591 Human Chromosome 7 STS S
gb_pat:AX153505	-	19.00	96.68	5.0e+03	191 ! AX153505 Sequence 9 from Patent
gb_pat:HUMFMD267A	+	19.00	96.68	5.0e+03	191 ! L15402 Homo sapiens (clone MFD2
gb_sts:G31733	+	19.00	96.68	5.0e+03	191 ! G31733 SWSS2689 Eric D. Green F
gb_ov:AB004302	+	19.00	96.35	5.2e+03	202 ! AB004302 Bos taurus mRNA for gr
gb_ov:HUMG6PDG06	+	19.00	96.32	5.3e+03	203 ! M5227 Human glucose-6-phosphat
gb_pl:AF250188	-	19.00	95.58	5.8e+03	230 ! AF250188 Zea mays transposon mu
gb_pat:AR112421	+	19.00	95.41	5.9e+03	237 ! AR112421 Sequence 11 from patent
gb_pl:CN501B2X	+	19.00	95.33	6.0e+03	240 ! AR113713 Botrytis cinerea strai
gb_ov:AF068239	+	19.00	95.21	6.1e+03	245 ! AF068239 Bacterium ClAb3 dichl
gb_sts:G59419	+	19.00	94.97	6.3e+03	255 ! GS9419 SHGC-130130 Human Homo s
gb_pr:HSU87908	-	19.00	94.93	6.3e+03	257 ! U87933 Human aconitase hydratase
gb_sts:G02071	-	19.00	94.81	6.4e+03	262 ! G02071 human STS 662AR, sequenc
gb_sts:G65760	-	19.00	94.81	6.4e+03	262 ! G65760 Dazl Miscellaneous mouse
gb_pr:HS52G1R	+	19.00	94.79	6.4e+03	263 ! Z61462 H.sapiens Cpg island DNA
gb_om:HLU03593	-	19.00	94.77	6.4e+03	264 ! U03593 Hydrurga leptonyx histon
gb_om:LUW03595	-	19.00	94.77	6.4e+03	264 ! U03595 Leptonychotes weddelli

gb_om:MAU03600	-	19.00	94.77	6.4e+03	264 ! U03600 Mirounga angustirostr
gb_om:PVU03589	-	19.00	94.77	6.4e+03	264 ! U03589 Phoca vitulina histon
gb_pr:HS121GIF	-	19.00	94.75	6.4e+03	265 ! Z54496 H.sapiens Cpg island
gb_ro:MUSIGFBP02	-	19.00	94.68	6.5e+03	268 ! L05437 Mouse insulin-like gr
gb_pr:HS121GIR	+	19.00	94.66	6.5e+03	269 ! Z54497 H.sapiens Cpg island
gb_pr:HSACO2G08	-	19.00	94.64	6.5e+03	270 ! AF093087 Homo sapiens aconit
gb_pat:I62369	-	19.00	94.59	6.6e+03	272 ! I62369 Sequence 10 from pate
gb_om:MIRHIS2AFP	-	19.00	94.55	6.6e+03	274 ! L17502 Mirounga leonina hist
gb_sts:G84229	-	19.00	94.53	6.6e+03	275 ! G84229 P288H2/T7 Human Chrom
gb_pat:AR084752	-	19.00	94.51	6.6e+03	276 ! AR084752 Sequence 78 from pa

seq_name: gb_pat:AU9389

seq_documentation_block:

LOCUS AU9389

DEFINITION oligonucleotide.

ACCESSION AU9389

VERSION AU9389.1 GI:513573

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 28)

AUTHORS

TITLE BINDING DOMAINS

JOURNAL Patent: WO 9201787-A 7 06-FEB-1992;

FEATURES

Location/Qualifiers

source

1..28

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 5 a 8 c 9 g 6 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AU9389 ..

Align seg 1/1 to: AU9389 from: 1 to: 28

1 GluAlaGly*****Ser 6

|||||

3 GAGCTGGCGCTCCAGT 20

seq_name: gb_pat:AR123943

seq_documentation_block:

LOCUS AR123943

DEFINITION Sequence 28 from patent US 6171823.

ACCESSION AR123943

VERSION AR123943.1 GI:14109304

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 32)

AUTHORS Woldike,H.Fabricius and Hastrup,S.

TITLE Process for producing extracellular proteins in bacteria

JOURNAL Patent: US 6171823-A 28 09-JAN-2001;

FEATURES

Location/Qualifiers

1..32

/organism="unknown"

BASE COUNT 6 a 15 c 8 g 3 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AR123943/rev ..

Align seg 1/1 to reverse of: AR123943 from: 1 to: 32

1 GluAlaGly*****Ser 6

|||||

30 GAGCGCGGGCGGCGAGC 13

seq_name: gb_pat:AR009555

seq_documentation_block:

LOCUS AR009555 34 bp DNA linear PAT 04-DEC-1998

DEFINITION Sequence 20 from patent US 5756323.

ACCESSION AR009555

VERSION AR009555.1 GI:3968360

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 34)

AUTHORS Kallenbach,S., Doyen,N. and Rougeon,F.

TITLE Method for generating structural and functional diversity in a

peptide sequence

JOURNAL Patent: US 5756323-A 20 26-MAY-1998;

FEATURES

Location/Qualifiers

1..34

source

BASE COUNT 6 a 10 c 12 g 6 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AR009555/rev ..

Align seg 1/1 to reverse of: AR009555 from: 1 to: 34

1 GluAlaGly*****Ser 6

|||||

28 GAGCGCGGATCCACTAGT 11

seq_name: gb_pat:AR009593

seq_documentation_block:

LOCUS AR009593 34 bp DNA linear PAT 04-DEC-1998

DEFINITION Sequence 58 from patent US 5756323.

ACCESSION AR009593

VERSION AR009593.1 GI:3968398

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 34)

AUTHORS Kallenbach,S., Doyen,N. and Rougeon,F.

TITLE Method for generating structural and functional diversity in a

peptide sequence

JOURNAL Patent: US 5756323-A 58 26-MAY-1998;

FEATURES

Location/Qualifiers

1..34

source

BASE COUNT 6 a 10 c 12 g 6 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AR009593/rev ..

Align seg 1/1 to reverse of: AR009593 from: 1 to: 34

1 GluAlaGly*****Ser 6

|||||

28 GAGCGCGGATCCACTAGT 11

seq_name: gb_pat:A19413

seq_documentation_block:

LOCUS A19413 36 bp DNA linear PAT 08-JUN-1994

DEFINITION oligonucleotide.

ACCESSION A19413

VERSION A19413.1 GI:583168

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 36)

AUTHORS van Gorcom,R.F.M., van Hartingsveldt,W., van Paridon,P.A.,

Veenstra,A.E., Luiten,R.G.M. and Selden,G.C.M.

TITLE Cloning and expression of microbial phytase

JOURNAL Patent: EP 0420358-A 2 03-APR-1991;

FEATURES

Location/Qualifiers

1..36

source

BASE COUNT 7 a 12 c 13 g 4 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x A19413 ..

Align seg 1/1 to: A19413 from: 1 to: 36

1 GluAlaGly*****Ser 6

|||||

2 .GAGCGGGGACTGCCAGT 19

seq_name: gb_pat:A19414

seq_documentation_block:

LOCUS A19414 36 bp DNA linear PAT 08-JUN-1994

DEFINITION oligonucleotide.

ACCESSION A19414

VERSION A19414.1 GI:583169

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 36)

AUTHORS van Gorcom,R.F.M., van Hartingsveldt,W., van Paridon,P.A.,

Veenstra,A.E., Luiten,R.G.M. and Selden,G.C.M.

TITLE Cloning and expression of microbial phytase

JOURNAL Patent: EP 0420358-A 3 03-APR-1991;

FEATURES

Location/Qualifiers

1..36

source

BASE COUNT 4 a 13 c 12 g 7 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000


```
ORIGIN
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM4 x A19414/rev ..
  Align seg 1/1 to reverse of: A19414 from: 1 to: 36
    1 GluAlaGly*****Ser 6
      |||||||
      35 GAGCGGGGACTGCCAGT 18
  seq_name: gb_pat:I13438
seq_documentation_block:
  LOCUS I13438 36 bp DNA linear PAT 26-JUL-1995
  DEFINITION Sequence 41 from patent US 5436156.
  ACCESSION I13438
  VERSION I13438.1 GI:910779
  KEYWORDS
  SOURCE Unknown.
  ORGANISM Unknown.
  UNCLASSIFIED.
  REFERENCE 1 (bases 1 to 36)
  AUTHORS Van Gorcom,R.F.M., Van Hartingsveldt,W., Van Paridon,P.A.,
  Veenstra,A.E., Luiten,R.G.M. and Selten,G.C.M.
  TITLE Cloning and expression of phytase from aspergillus
  JOURNAL Patent: US 5436156-A 41 25-JUL-1995;
  FEATURES
    source
      1..36
      /organism="unknown"
BASE COUNT 7 a 12 c 13 g 4 t
ORIGIN
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM4 x I13438 ..
  Align seg 1/1 to: I13438 from: 1 to: 36
    1 GluAlaGly*****Ser 6
      |||||||
      2 GAGCGGGGACTGCCAGT 19
  seq_name: gb_pat:I13439
seq_documentation_block:
  LOCUS I13439 36 bp DNA linear PAT 26-JUL-1995
  DEFINITION Sequence 42 from patent US 5436156.
  ACCESSION I13439
  VERSION I13439.1 GI:910780
  KEYWORDS
  SOURCE Unknown.
  ORGANISM Unknown.
  REFERENCE 1 (bases 1 to 36)
  AUTHORS Van Gorcom,R.F.M., Van Hartingsveldt,W., Van Paridon,P.A.,
  Veenstra,A.E., Luiten,R.G.M. and Selten,G.C.M.
  TITLE Cloning and expression of phytase from aspergillus
  JOURNAL Patent: US 5436156-A 42 25-JUL-1995;
  FEATURES
    source
      1..36
      /organism="unknown"
BASE COUNT 7 a 12 c 13 g 4 t
ORIGIN
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM4 x I13438 ..
  Align seg 1/1 to: I13438 from: 1 to: 36
    1 GluAlaGly*****Ser 6
      |||||||
      2 GAGCGGGGACTGCCAGT 19
  seq_name: gb_pr:HSU30392
seq_documentation_block:
  LOCUS HSU30392 42 bp mRNA linear PRI 01-AUG-1995
  DEFINITION Human isolate M103 T-cell receptor alpha V-J junction (TCR Valpha
  3/J alpha 44) mRNA, partial cds.
  ACCESSION U30392
  VERSION U30392.1 GI:915419
  KEYWORDS
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 42)
  AUTHORS Dave,V.P., Larche,M., Rencher,S.D., Koop,B.F. and Hurwitz,J.L.
  TITLE Restricted usage of T-cell receptor V alpha sequence and
  variable-jointing pairs after normal T-cell development and bone
  marrow transplantation
  JOURNAL Hum. Immunol. 37 (3), 178-184 (1993)
  MEDLINE 94064390
  REFERENCE 2 (bases 1 to 42)
  AUTHORS Hurwitz,J.L.
  TITLE Direct Submission
  JOURNAL Submitted (22-JUN-1995) Julia L. Hurwitz, St. Jude Children's
  Research Hospital, Immunology, 332 N. Lauderdale, Memphis, TN
  38101, USA
  FEATURES
    Location/Qualifiers
      source
        1..42
        /organism="Homo sapiens"
        /isolate="M103"
        /db_xref="taxon:9606"
        /tissue_type="blood"
        1..42
        <1..>42
        /gene="TCR Valpha 3/J alpha 44"
        /codon_start=1
        /gene="TCR Valpha 3/J alpha 44"
        /product="T-cell receptor alpha V-J junction"
        /protein_id="AAA73581.1"
        /db_xref="GI:915420"
        /translation="FCATEAGTASKLTF"
BASE COUNT 9 a 13 c 9 g 11 t
ORIGIN
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM4 x HSU30392 ..
  Align seg 1/1 to: HSU30392 from: 1 to: 42
```

```
1 GluAlaGly*****Ser 6
|||||
13 GAAGCGGCGCTCCAGT 30

seq_name: gb_pat:AX162903

seq_documentation_block:
LOCUS AX162903 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6231 from Patent WO0140521.
ACCESSION AX162903
VERSION AX162903.1 GI:14544234
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shinkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0140521-A 6231 07-JUN-2001;
Curagen Corporation (US)
FEATURES
Source 1..51
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 26
/note="1 of 2 allelic variants (6232 is other entry)"
BASE COUNT 9 a 13 c 17 g 12 t
ORIGIN
alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

seq_documentation_block:
LOCUS AX162903 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6232 from Patent WO0140521.
ACCESSION AX162903
VERSION AX162903.1 GI:14544235
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shinkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0140521-A 6232 07-JUN-2001;
Curagen Corporation (US)
FEATURES
Source 1..51
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 26
/note="2 of 2 allelic variants (6231 is other entry)"
BASE COUNT 9 a 12 c 17 g 13 t
ORIGIN

1 GluAlaGly*****Ser 6
|||||
29 GAAGCTGGTCCACCACT 46

seq_name: gb_pat:AX162904

seq_documentation_block:
LOCUS AX162904 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6233 from Patent WO0140521.
ACCESSION AX162904
VERSION AX162904.1 GI:14544236
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shinkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0140521-A 6233 07-JUN-2001;
Curagen Corporation (US)
FEATURES
Source 1..51
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 26
/note="1 of 2 allelic variants (6234 is other entry)"
BASE COUNT 9 a 12 c 18 g 12 t
ORIGIN

1 GluAlaGly*****Ser 6
|||||
24 GAAGCTGGTCCACCACT 41

seq_name: gb_pat:AX162906

seq_documentation_block:
LOCUS AX162906 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6234 from Patent WO0140521.
ACCESSION AX162906
VERSION AX162906.1 GI:14544237
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
```

AUTHORS Shimkets,R.A. and Leach,M.
 TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
 JOURNAL Patent: WO 0140521-A 6234 07-JUN-2001.;
 Curagen Corporation (US)
 FEATURES
 source 1. .51
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 misc_feature 26
 /note="2 of 2 allelic variants (6233 is other entry)
 Accession number cg44917278"
 BASE COUNT 8 a 12 c 19 g 12 t
 ORIGIN
 alignment_scores
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 BASK-853-CLAIM4 x AX162906 ..
 Align seg 1/1 to: AX162906 from: 1 to: 51
 1 GluAlaGly*****Ser 6
 |||||
 24 GAGGCTGGTTCACCAAGT 41
 seq_name: gb_pr:AF011601
 seq_documentation_block:
 LOCUS AF011601 51 bp mRNA linear PRI 21-JAN-1998
 DEFINITION Homo sapiens T cell receptor beta chain (BV6S7*2-BJ2S7) mRNA,
 partial cds.
 ACCESSION AF011601
 VERSION AF011601.1 GI:2654670
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 51)
 AUTHORS Rowen,L., Koop,B.F. and Hood,L.
 TITLE The complete 685-kilobase DNA sequence of the human beta T cell
 receptor locus
 JOURNAL Science 272 (5269), 1755-1762 (1996)
 MEDLINE 96256474
 REFERENCE 2 (bases 1 to 51)
 AUTHORS Ciurli,C., Posnett,D.N., Sekaly,R.P. and Denis,F.
 TITLE Highly biased CDR3 usage in restricted sets of beta chain variable
 regions during viral superantigen 9 response
 JOURNAL J. Exp. Med. 187 (2), 253-258 (1998)
 MEDLINE 98095705
 REFERENCE 3 (bases 1 to 51)
 AUTHORS Ciurli,C., Sekaly,R.-P. and Denis,F.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-1997) Immunology, IRCM, 110 Pine Avenue West,
 Montreal, Quebec H2W 1R7, Canada
 FEATURES
 source 1. .51
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="28S"
 <1. .>51
 /gene="BV6S7*2-BJ2S7"
 misc_feature 1. .51
 /gene="BV6S7*2-BJ2S7"
 /note="T cell receptor functional region"
 <1. .>51
 /gene="BV6S7*2-BJ2S7"
 CDS
 /codon_start=1

/product="T cell receptor beta chain"
 /protein_id="AAC51967.1"
 /db_xref="GI:2654671"
 /translation="CASSFGSGGYFGPGTR"
 BASE COUNT 8 a 14 c 22 g 7 t
 ORIGIN
 alignment_scores
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 BASK-853-CLAIM4 x AF011601 ..
 Align seg 1/1 to: AF011601 from: 1 to: 51
 1 GluAlaGly*****Ser 6
 |||||
 20 GAAGCGGGCAGTACTTCG 37
 seq_name: gb_pat:AR053957
 seq_documentation_block:
 LOCUS AR053957 84 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 94 from patent US 5834286.
 ACCESSION AR053957
 VERSION AR053957.1 GI:5978819
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 84)
 AUTHORS Niettinen,H.K.M., Paloheimo,M.T., Fagerstrom,R.B.,
 Miettinen-Oinonen,A.S.K., Turunen,M.K., Rambosek,J.A.,
 Piddington,C.S., Houston,C.S. and Cantrell,M.A.
 TITLE Recombinant cells that express phytate degrading enzymes in desired
 ratios
 JOURNAL Patent: US 5834286-A 94 10-NOV-1998;
 FEATURES Location/Qualifiers
 source 1. .84
 BASE COUNT 18 a 21 c 32 g 13 t
 ORIGIN
 alignment_scores
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 BASK-853-CLAIM4 x AR053957 ..
 Align seg 1/1 to: AR053957 from: 1 to: 84
 1 GluAlaGly*****Ser 6
 |||||
 3 GAGCGGGGACTGCCAGT 20
 seq_name: gb_in:AF185208
 seq_documentation_block:
 LOCUS AF185208 133 bp DNA linear INV 20-OCT-1999
 DEFINITION Pista australis U2 small nuclear RNA, partial sequence.
 ACCESSION AF185208
 VERSION AF185208.1 GI:6049713
 KEYWORDS
 SOURCE Pista australis.
 ORGANISM Pista australis
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
 Terebellida; Terebellidae; Pista.

```

REFERENCE 1 (bases 1 to 133)
AUTHORS Brown,S., Rouse,G., Hutchings,P. and Colgan,D.
TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in
        analyses of polychaete relationships
JOURNAL Aust. J. Zool. 47 (1999) In press
REFERENCE 2 (bases 1 to 133)
AUTHORS Brown,S., Rouse,G., Hutchings,P. and Colgan,D.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6
        College Street, Sydney, NSW 2000, Australia
FEATURES   Location/Qualifiers
            source
                1..133
                    /organism="Pista australis"
                    /specimen_voucher="AMW24390"
                    /db_xref="taxon:104725"
                    <1..>133
                    /note="spliceosome component"
                    /product="U2 small nuclear RNA"
BASE COUNT 29 a 29 c 32 g 43 t
ORIGIN

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x AF185208/rev ..
Align seg 1/1 to reverse of: AF185208 from: 1 to: 133
1 GluAlaGly*****Ser 6
|||||
119 GAAGCGGAGCAAGCTC 102

seq_name: gb_vi:HIV8417

seq_documentation_block:
LOCUS HIV8417 141 bp DNA linear VRL 29-MAY-1998
DEFINITION Human immunodeficiency virus type 2 LTR, isolate b1241.
ACCESSION AJ008417
VERSION AJ008417.1 GI:3171849
KEYWORDS LTR.
SOURCE Human immunodeficiency virus type 2.
ORGANISM Human immunodeficiency virus type 2
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE 1 (bases 1 to 141)
AUTHORS Xiang,Z., Ariyoshi,K., Wilkins,A., Dias,F., Whittle,H. and
        Breuer,J.
TITLE HIV type 2 pathogenicity is not related to subtype in rural Guinea
        Bissau
JOURNAL AIDS Res. Hum. Retroviruses 13 (6), 501-505 (1997)
MEDLINE 97255645
AUTHORS Grassly,N., Xiang,Z., Ariyoshi,K., Aaby,P., Jensen,H., Dias,F.,
        Whittle,H. and Breuer,J.
TITLE Correlation of HIV-2 genotype with progression to AIDS in vivo
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 141)
AUTHORS Xiang,Z.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1998) Xiang Z., London Hospital Medical College,
        Department of Virology, Queen Mary Westfield, Turner Street,
        London, E1 2AD, UK
FEATURES   Location/Qualifiers
            source
                1..141
                    /organism="Human immunodeficiency virus type 2"
                    /strain="A"
                    /isolate="b1241"
                    /db_xref="taxon:11709"
                    1..141
LTR

REFERENCE 1 (bases 1 to 141)
AUTHORS Brown,S., Rouse,G., Hutchings,P. and Colgan,D.
TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in
        analyses of polychaete relationships
JOURNAL Aust. J. Zool. 47 (1999) In press
REFERENCE 2 (bases 1 to 141)
AUTHORS Brown,S., Rouse,G., Hutchings,P. and Colgan,D.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6
        College Street, Sydney, NSW 2000, Australia
FEATURES   Location/Qualifiers
            source
                1..141
                    /organism="Human immunodeficiency virus type 2"
                    /strain="A"
                    /isolate="b1241"
                    /db_xref="taxon:11709"
                    1..141
LTR

```

```

BASE COUNT 25 a 35 c 45 g 36 t
ORIGIN

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x HIV8417 ..
Align seg 1/1 to: HIV8417 from: 1 to: 141
1 GluAlaGly*****Ser 6
|||||
79 GAGGCTGGCAGTTCGAGC 96

seq_name: gb_vi:HIV8341

```

```

seq_documentation_block:
LOCUS HIV8341 147 bp DNA linear VRL 29-MAY-1998
DEFINITION Human immunodeficiency virus type 2 LTR, isolate b1046.
ACCESSION AJ008341
VERSION AJ008341.1 GI:3171774
KEYWORDS LTR.
SOURCE Human immunodeficiency virus type 2.
ORGANISM Human immunodeficiency virus type 2
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE 1 (bases 1 to 147)
AUTHORS Xiang,Z., Ariyoshi,K., Wilkins,A., Dias,F., Whittle,H. and
        Breuer,J.
TITLE HIV type 2 pathogenicity is not related to subtype in rural Guinea
        Bissau
JOURNAL AIDS Res. Hum. Retroviruses 13 (6), 501-505 (1997)
MEDLINE 97255645
AUTHORS Grassly,N., Xiang,Z., Ariyoshi,K., Aaby,P., Jensen,H., Dias,F.,
        Whittle,H. and Breuer,J.
TITLE Correlation of HIV-2 genotype with progression to AIDS in vivo
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 147)
AUTHORS Xiang,Z.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1998) Xiang Z., London Hospital Medical College,
        Department of Virology, Queen Mary Westfield, Turner Street,
        London, E1 2AD, UK
FEATURES   Location/Qualifiers
            source
                1..147
                    /organism="Human immunodeficiency virus type 2"
                    /strain="A"
                    /isolate="b1046"
                    /db_xref="taxon:11709"
                    1..147
LTR

BASE COUNT 29 a 37 c 44 g 37 t
ORIGIN

```

```

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x HIV8341 ..
Align seg 1/1 to: HIV8341 from: 1 to: 147
1 GluAlaGly*****Ser 6
|||||
82 GAGGCTGGCAGTTCGAGC 99

```

```

seq_name: gb_ov:CLDABC
seq_documentation_block:
LOCUS CLDABC 156 bp DNA linear VRT 22-APR-1997
DEFINITION C.livia gene encoding alpha-B-crystallin.
ACCESSION X96594
VERSION X96594.1 GI:1945728
KEYWORDS alpha-B-crystallin.
SOURCE domestic pigeon.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Columbiformes; Columba.
REFERENCE
1 (bases 1 to 156)
AUTHORS Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
TITLE alpha-Crystallin sequences support a galliform/anseriform clade
JOURNAL Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
MEDLINE 97271643
REFERENCE
2 (bases 1 to 156)
AUTHORS Caspers,G.J.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
NETHERLANDS
FEATURES
source Location/Qualifiers
1..156
/organism="Columba livia"
/db_xref="taxon:8932"
<1..>156
/number=1
<1..>156
/codon_start=1
/product="alpha-B-crystallin"
/protein_id="CAA65412.1"
/db_xref="GI:1945729"
/db_xref="SWISS-PROT:O12987"
/translation="LIRRLSLWLPASRFIDQIFGHELOESELLPASPSFPLMRSP
ILRPSWL"
BASE COUNT 27 a 50 c 36 g 43 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x CLDABC/rev ..
Align seg 1/1 to reverse of: CLDABC from: 1 to: 156

1 GluAlaGly*****Ser 6
|||||
98 GAAGCGGGGAGCAGCTCT 81

seq_name: gb_ov:EERABC
seq_documentation_block:
LOCUS EERABC 156 bp mRNA linear VRT 22-APR-1997
DEFINITION E.elegans mRNA for alpha-B-crystallin.
ACCESSION X96596
VERSION X96596.1 GI:1945737
KEYWORDS alpha-B-crystallin.
SOURCE elegant crested-tinamou.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Palaeognathae; Tinamiformes; Tinamidae;
Eudromia.
REFERENCE
1 (bases 1 to 156)
AUTHORS Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
TITLE alpha-Crystallin sequences support a galliform/anseriform clade
JOURNAL Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
MEDLINE 97271643

```

```

seq_name: gb_ov:CLDABC
seq_documentation_block:
LOCUS CLDABC 156 bp DNA linear VRT 22-APR-1997
DEFINITION C.livia gene encoding alpha-B-crystallin.
ACCESSION X96594
VERSION X96594.1 GI:1945728
KEYWORDS alpha-B-crystallin.
SOURCE domestic pigeon.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Columbiformes; Columba.
REFERENCE
1 (bases 1 to 156)
AUTHORS Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
TITLE alpha-Crystallin sequences support a galliform/anseriform clade
JOURNAL Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
MEDLINE 97271643
REFERENCE
2 (bases 1 to 156)
AUTHORS Caspers,G.J.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
NETHERLANDS
FEATURES
source Location/Qualifiers
1..156
/organism="Columba livia"
/db_xref="taxon:8932"
<1..>156
/number=1
<1..>156
/codon_start=1
/product="alpha-B-crystallin"
/protein_id="CAA65412.1"
/db_xref="GI:1945729"
/db_xref="SWISS-PROT:O12987"
/translation="LIRRLSLWLPASRFIDQIFGHELOESELLPASPSFPLMRSP
ILRPSWL"
BASE COUNT 27 a 50 c 36 g 43 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x CLDABC/rev ..
Align seg 1/1 to reverse of: CLDABC from: 1 to: 156

1 GluAlaGly*****Ser 6
|||||
98 GAAGCGGGGAGCAGCTCT 81

seq_name: gb_ov:EERABC
seq_documentation_block:
LOCUS EERABC 156 bp mRNA linear VRT 22-APR-1997
DEFINITION E.elegans mRNA for alpha-B-crystallin.
ACCESSION X96596
VERSION X96596.1 GI:1945737
KEYWORDS alpha-B-crystallin.
SOURCE elegant crested-tinamou.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Palaeognathae; Tinamiformes; Tinamidae;
Eudromia.
REFERENCE
1 (bases 1 to 156)
AUTHORS Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
TITLE alpha-Crystallin sequences support a galliform/anseriform clade
JOURNAL Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
MEDLINE 97271643

```

```

REFERENCE 2 (bases 1 to 156)
AUTHORS Caspers,G.J.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
NETHERLANDS
FEATURES
source Location/Qualifiers
1..156
/organism="Eudromia elegans"
/db_xref="taxon:8805"
<1..>156
/codon_start=1
/product="alpha-B-crystallin"
/protein_id="CAA65414.1"
/db_xref="GI:1945738"
/db_xref="SWISS-PROT:O12991"
/translation="LIRRLSLWLPASRFIDQIFGHELOESELLPASPGFSFLMRSP
IFRPSWV"
BASE COUNT 20 a 56 c 41 g 39 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x EERABC/rev ..
Align seg 1/1 to reverse of: EERABC from: 1 to: 156

1 GluAlaGly*****Ser 6
|||||
98 GAGCGGGGAGCAGCTCT 81

seq_name: gb_ov:RNUT78130
seq_documentation_block:
LOCUS RNUT78130 157 bp mRNA linear ROD 27-NOV-1996
DEFINITION Rattus norvegicus zinc finger protein 2 (DZF2) mRNA, partial cds.
ACCESSION U78130
VERSION U78130.1 GI:1685325
KEYWORDS
SOURCE Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
REFERENCE
1 (bases 1 to 157)
AUTHORS Gebelein,B., Mesa,K. and Urrutia,R.
TITLE A novel profile of expressed sequence tags for zinc finger encoding
genes from the poorly differentiated exocrine pancreatic cell line
AR4IP
JOURNAL Cancer Lett. 105 (2), 225-231 (1996)
MEDLINE 96319713
REFERENCE
2 (bases 1 to 157)
AUTHORS Mesa,K., Gebelein,B., Cook,T. and Urrutia,R.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-1996) GI Research Unit, Mayo Clinic, 200 First
Street SW, Rochester, MN 55905, USA
FEATURES
source Location/Qualifiers
1..157
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_line="AR4IP"
1..157
/gene="DZF2"
<1..>157
/gene="DZF2"
/product="zinc finger protein 2"
/protein_id="AAB36802.1"

```

```

/db_xref="GI:1685326"
/translation="RVGKSRLLSQITQHSIHTGKPECKEKGTFRLYFSLSQHQ
RIHTGSKP"
BASE COUNT      40 a      46 c      38 g      33 t
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x RNU78130/rev ..
Align seg 1/1 to reverse of: RNU78130 from: 1 to: 157

1 GluAlaGly*****Ser 6
|||||
25 GAGCGCGGACGACTTCC 8

seq_name: gb_ro:MWJ004556

seq_documentation_block:
LOCUS      MWJ004556      162 bp      DNA      linear      ROD 21-OCT-1999
DEFINITION Maxomys whiteheadi M. whi741-17 DNA for L1 (LINE-1) DNA element.
ACCESSION  AJ004556
VERSION    AJ004556.1 GI:6102990
KEYWORDS   LINE element.
SOURCE     Whitehead's spiny rat.
ORGANISM   Maxomys whiteheadi
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Maxomys.
REFERENCE  1 (bases 1 to 162)
AUTHORS   Verneau,O., Catzefflis,F. and Furano,A.V.
TITLE     Determining and dating recent rodent speciation events by using L1
            (LINE-1) retrotransposons
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11284-11289 (1998)
MEDLINE    98409649
REFERENCE  2 (bases 1 to 162)
AUTHORS   Verneau,O.
TITLE     Direct Submission
JOURNAL    Submitted (15-Oct-1997) Verneau O., Laboratoire de Biologie
            Animale, Centre de Biologie et d'Ecologie, Tropicale et
            Meditorananne, Universite de Perpignan, F-66860 Perpignan, France
FEATURES   Location/Qualifiers
            1..162
            /organism="Maxomys whiteheadi"
            /db_xref="taxon:69127"
            /clone_lib="M. whi741-17"
            <1..>162
            /note="LINE-1 (Long Interspersed Nuclear Element)"
BASE COUNT      53 a      39 c      37 g      33 t
ORIGIN

misc_feature
            1..162
            /organism="Maxomys whiteheadi"
            /db_xref="taxon:69127"
            /clone_lib="M. whi741-17"
            <1..>162
            /note="LINE-1 (Long Interspersed Nuclear Element)"
BASE COUNT      53 a      39 c      37 g      33 t
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x MWJ004556 ..
Align seg 1/1 to: MWJ004556 from: 1 to: 162

1 GluAlaGly*****Ser 6
|||||
119 GAGCGCGGTGCTAGCAGC 136

seq_name: gb_sts:DM62C3T

```

```

seq_documentation_block:
LOCUS      DM62C3T      166 bp      DNA      linear      STS 17-APR-1996
DEFINITION D. melanogaster STS determined from European Mapping Project
            cosmid, sequence tagged site.
ACCESSION  Z32221
VERSION    Z32221.1 GI:1150240
KEYWORDS   STS.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 166)
TITLE     European Drosophila Mapping Consortium.
JOURNAL    Direct Submission
            Submitted (08-APR-1994) Michael Ashburner, Department of Genetics,
            Downing St., Cambridge CB2 3EH, England
REFERENCE  2 (bases 1 to 166)
AUTHORS   European Drosophila Mapping Consortium.
TITLE     Direct Submission
JOURNAL    Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,
            Downing St., Cambridge CB2 3EH, England
REMARK     Updated comments
REFERENCE  3 (bases 1 to 166)
AUTHORS   Madueno,E., Papagiannakis,G., Rimmington,G.A., Saunders,R.D.C.,
            Savakis,C., Siden-Kiamos,I., Skavdis,G., Spanos,L., Trencear,J.,
            Adam,P., Ashburner,M., Benos,P., Bolshakov,V.N., Coulson,D.,
            Glover,D.M., Herrmann,S., Kafatos,F.C., Louis,C., Majerus,T. and
            Modolell,J.
TITLE     A physical map of the X chromosome of Drosophila melanogaster:
            cosmid contigs and sequence tagged sites
JOURNAL    Genetics 139 (4), 1631-1647 (1995)
MEDLINE    95309678
COMMENT     On Jan 7, 1996 this sequence version replaced gi:938135.
            STS_name = Dm62C3T
            clone_name = 62C3
            STS_from_promoter = T7
            vector_class = cosmid, Lorist 6
            origin_of_clone = Oregon-R
            in_situ_site_primary = 2F-3A
            STS_dbSTS_AC = 4773
            BLAST_program = BLASTN
            database_searched = EMBL
            database_version = 45.0 and updates till date_of_search
            date_of_search = 08-01-1996
            FlyBase_gene_symbol = ph-d
            FlyBase_gene_name = polyhomeotic distal
            FlyBase_gene_identifier = FBgn0004860
            AC_number_of_hit = M64750
            probability_of_hit = 1.00E-56
            HSP_score_of_hit = 748
            BLAST_program = BLASTX
            database_searched = SWISSPROT
            database_version = 32.0
            date_of_search = 15-12-1995.
FEATURES   Location/Qualifiers
            1..166
            /organism="Drosophila melanogaster"
            /strain="Oregon-R"
            /db_xref="taxon:7227"
            /clone="62C3"
BASE COUNT      31 a      51 c      35 g      42 t      7 others
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x DM62C3T/rev ..

```

Align seg 1/1 to reverse of: DM62C3T from: 1 to: 166

1 GluAlaGly*****Ser 6
|||||
87 GAAGCTGGGTCCAGCTCT 70

seq_name: gb_sts:HUMSWS229

seq_documentation_block:

LOCUS HUMSWS229 171 bp DNA linear STS 02-MAY-2000
DEFINITION Human Chromosome 7 STS swSS229; single read, sequence tagged site.

ACCESSION L10591

VERSION L10591.1 GI:484308

KEYWORDS STS; STS sequence; primer; sequence tagged site.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 171)

AUTHORS Green, E.D., Mohr, R.M., Idol, J.R., Jones, M., Buckingham, J.M.,

Deaven, L.L., Moyzis, R.K. and Olson, M.V.

TITLE Systematic generation of sequence-tagged sites for physical mapping
of human chromosomes: application to the mapping of human

chromosome 7 using yeast artificial chromosomes

Genomics 11, 548-564 (1991)

JOURNAL MEDLINE 92128937

REFERENCE 2 (bases 1 to 171)

AUTHORS Green, E.D.

TITLE Direct Submission

Submitted (02-FEB-1993); Eric D. Green, Center for Genetics in

Medicine, Washington University School of Medicine, 4566 Scott

Avenue, St. Louis, MO 63110, USA

On May 11, 1994 this sequence version replaced gi:292580.

e-mail: egreen@wugenmail.wustl.edu

Primer A: AATCAGCATCTCTGGTG

Primer B: CTGTTCCCTGAAACTTCTGTG

STS Size: 60 bp

Template: 30-100 ng

Primer: 1 uM

dNTPs: 200 uM

MgCl2: 1.5 mM

KCl: 50 mM

TrisHCl: 10 mM

Taq Polymerase: 0.05 units/ul

pH: 8.3

Total vol: 5 ul

PCR Profile: Denaturation: 92 degrees C for 1.00
minute(s)
Annealing: 60 degrees C for 2.00 minute(s)
Polymerization: 72 degrees C for 2.00 minute(s)

PCR Cycles: 35

Thermocycler: PE TC.

Location/Qualifiers

1. .171

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="7"

/standard_name="swSS229"

106. .165

106. .123

primer_bind

primer_bind

complement(145. .165)

38 a 54 c 43 g 35 t 1 others

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x HUMSWS229/rev

..

Align seg 1/1 to reverse of: HUMSWS229 from: 1 to: 171

1 GluAlaGly*****Ser 6
|||||
29 GAGGCGAGTGTGCGTCC 12

seq_name: gb_pat:AX195305

seq_documentation_block:

LOCUS AX195305 191 bp DNA linear PAT 28-AUG-2001

DEFINITION Sequence 9 from Patent WO0151631.

ACCESSION AX195305

VERSION AX195305.1 GI:15385854

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 191)

AUTHORS Reske-Kunz, A., Ross, X., Ross, R. and Bros, M.

TITLE Regulatory sequence for the specific expression in dendritic cells
and uses thereof

JOURNAL Patent: WO 0151631-A 9 19-JUL-2001;

Reske-Kunz, Angelika (DE) ; Ross, Xiaolan (DE) ; Ross, Ralf (DE) ;

Bros, Matthias (DE)

Location/Qualifiers

1. .191

/organism="Homo sapiens"

/db_xref="taxon:9606"

19 a 82 c 62 g 28 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AX195305/rev

..

Align seg 1/1 to reverse of: AX195305 from: 1 to: 191

1 GluAlaGly*****Ser 6

|||||

26 GAGGCGGCGGCGAGCTCC 9

seq_name: gb_pr:HUMMFD267A

seq_documentation_block:

LOCUS HUMMFD267A 191 bp DNA linear PRI 13-JUL-1993

DEFINITION Homo sapiens (clone MFD267) PCR primer.

ACCESSION L15402

VERSION L15402.1 GI:292304

KEYWORDS PCR; PCR amplified; PCR identification; PCR primer.

SOURCE Homo sapiens blood DNA.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Weber, J.L.

JOURNAL Unpublished (1993)

REFERENCE 2 (bases 1 to 191)

AUTHORS Weber, J.L., Wang, Z., Hansen, K., Stephenson, M., Kappel, C.,

Salzman, S., Wilkie, P.J., Keats, B., Dracopoli, N.C., Brandriff, B.F.

and Olsen, A.S.

Evidence for human meiotic crossover interference obtained through

construction of a short tandem repeat polymorphism linkage map of

chromosome 19

Unpublished (1993)

JOURNAL

FEATURES

Location/Qualifiers

1. .191

misc_feature
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="blood"
1..191

primer_bind
/standard_name="Only one strand sequenced"
/evidence-experimental
28..47

repeat_region
/evidence-experimental
69..117

primer_bind
/rpt_family="(TGTA)n"
/rpt_type=tandem
/evidence-experimental
complement(114..133)
/standard_name="PCR primer"
/evidence-experimental
45 a 30 c 42 g 74 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x HUMMFD267A ..

Align seg 1/1 to: HUMMFD267A from: 1 to: 191

1 GluAlaGly*****Ser 6
|||||

125 GAGGAGGGTCTAGCTCT 142

seq_name: gb_sts:G31733

seq_documentation_block:
LOCUS G31733 191 bp DNA linear STS 28-SEP-1998

DEFINITION SW52689 Eric D. Green Homo sapiens STS genomic, sequence tagged

site.
G31733
VERSION G31733.1 GI:1916458

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
Leckie,M.P. and Green,E.D.

TITLE A collection of 1814 human chromosome 7-specific STSs
MEDLINE
PUBMED 97189344

REFERENCE
AUTHORS Green,E.D.
TITLE Human chromosome 7 STSs (1997)

COMMENT
Synonyms: MFD267
GDB: GDB:195021
GDB_DSEG: D7S558
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Primer A: GGAATCTGCTAGACTGTTT
Primer B: CCCTGCCCTCTAAATATAC
STS size: 106
PCR Profile:
Presoak: 0 degrees C for 0.00 minute(s)

Denaturation: 92 degrees C for 0.17 minute(s)
Annealing: 55 degrees C for 1.00 minute(s)
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600

Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH4Cl: 5 mM
pH: 8.6

This STS has been incorporated into the NHGRI chromosome 7
physical map, but was developed by another investigator. See
GenBank record: L15402 For additional information about the NHGRI
chromosome 7 mapping project, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7. Also see Genomics
11:548-64 (1991) [MUID=92128937].

FEATURES
source
1..191
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
clone_lib="Eric D. Green"
28..133
primer_bind 28..47
primer_bind complement(114..133) 74 t
BASE COUNT 45 a 30 c 42 g 74 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x G31733 ..

Align seg 1/1 to: G31733 from: 1 to: 191

1 GluAlaGly*****Ser 6
|||||

125 GAGGAGGGTCTAGCTCT 142

seq_name: gb_om:AB004302

seq_documentation_block:
LOCUS AB004302 202 bp mRNA linear MAM 25-MAR-1998

DEFINITION Bos taurus mRNA for growth differentiation factor-7, partial cds.

ACCESSION AB004302
VERSION AB004302.1 GI:2190035

KEYWORDS growth differentiation factor-7.
SOURCE Bos taurus CDNA to mRNA.
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS Oida,S.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-1997) Shinichiro Oida, Tokyo Medical and Dental
University, Faculty of Dentistry, Department of Oral Biochemistry;
Yushima 1-5-45, Bunkyo-ku, Tokyo 113, Japan
(E-mail:yoko-m.perident.tmd.ac.jp, Tel:+81-3-5803-5449)

REFERENCE
2 (sites)


```
AUTHORS Morotome,Y., Goseki-Sone,M., Ishikawa,I. and Oida,S.
TITLE Gene expression of growth and differentiation factors-5, -6, and -7
        in developing bovine tooth at the root forming stage
JOURNAL Biochem. Biophys. Res. Commun. 244 (1), 85-90 (1998)
MEDLINE 98183403
REMARK Erratum:[published erratum appears in Biochem Biophys Res Commun
1998 May 29;246(3):925]]
FEATURES
source      Location/Qualifiers
            1..202
            /organism="Bos taurus"
            /db_xref="taxon:9913"
            <1..>202
            /codon_start=2
            /product="growth differentiation factor-7"
            /protein_id="BAA20364.1"
            /db_xref="GI:2190036"
            /translation="PDYEAYHCEGVCDFFPLRSHLEPTNHAIQITLLNSMAPDAAPAS
            CCVPARLSPISILYIDAANNVY"
BASE COUNT 35 a 83 c 51 g 33 t
ORIGIN
alignment_scores
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block
BASK-853-CLAIM4 x AB004302/rev
Align seg 1/1 to reverse of: AB004302 from: 1 to: 202
1 GluAlaGly*****Ser 6
|||||
132 GAGGCTGGCGCGCATCC 115
seq_name: gb_pr:HUMG6PDG06
seq_documentation_block
LOCUS HUMG6PDG06 203 bp DNA linear PRI 22-MAR-2001
DEFINITION Human glucose-6-phosphate dehydrogenase gene, exon 6.
ACCESSION M65227
VERSION M65227.1 GI:182881
KEYWORDS glucose-6-phosphate dehydrogenase.
SEGMENT 6 of 13
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203)
AUTHORS Martini,G., Toniolo,D., Vulliamy,T., Luzzatto,L., Dono,R.,
Viglietto,G., Paonessa,G., D'Urso,M. and Persico,M.G.
TITLE Structural analysis of the X-linked gene encoding human glucose
6-phosphate dehydrogenase
JOURNAL EMBO J. 5 (8), 1849-1855 (1986)
MEDLINE 87004551
REFERENCE 2 (sites)
AUTHORS Hirono,A. and Beutler,E.
TITLE Molecular cloning and nucleotide sequence of cDNA for human
glucose-6-phosphate dehydrogenase variant A(-)
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (11), 3951-3954 (1988)
MEDLINE 88234536
FEATURES
source      Location/Qualifiers
            1..203
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="Xq28"
            /sex="female"
            /tissue_lib="Claron 4A"
            29..187
            /gene="G6PD"
            /note="G00-120-621"
            /number=6
exon
```

```
source      29..187
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="A-19"
            /haplotype="black"
            /sex="male"
            /cell_type="Epstein-Barr virus transformed lymphoblastoid
            cell"
BASE COUNT 42 a 63 c 64 g 34 t
ORIGIN
alignment_scores
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block
BASK-853-CLAIM4 x HUMG6PDG06
Align seg 1/1 to: HUMG6PDG06 from: 1 to: 203
1 GluAlaGly*****Ser 6
|||||
28 GAGGCTGGACCGCATCA 45
seq_name: gb_pl:AF250188
seq_documentation_block
LOCUS AF250188 230 bp DNA linear PLN 24-APR-2001
DEFINITION Zea mays transposon mutator insertion band 7, partial sequence.
ACCESSION AF250188
VERSION AF250188.1 GI:13774279
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 230)
AUTHORS Rabinowicz,P.D. and Grotewold,E.
TITLE A novel reverse-genetic approach (SIMF) identifies Mutator
insertions in new Myb genes
JOURNAL Planta 211 (6), 887-893 (2000)
MEDLINE 21017287
REFERENCE 2 (bases 1 to 230)
AUTHORS Rabinowicz,P.D. and Grotewold,E.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Cold Spring Harbor Laboratory, 1 Bungtown
Rd., Cold Spring Harbor, NY 11724, USA
FEATURES
source      Location/Qualifiers
            1..230
            /organism="Zea mays"
            /db_xref="taxon:4577"
            <1..17
            /organism="Zea mays"
            /transposon="mutator"
            /db_xref="taxon:4577"
            /note="mutator insertion band 7"
repeat_region <1..17
BASE COUNT 44 a 55 c 81 g 50 t
ORIGIN
alignment_scores
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block
BASK-853-CLAIM4 x AF250188/rev
```

Align seg 1/1 to reverse of: AF250188 from: 1 to: 230

1 GluAlaGly*****Ser 6
|||||
65 GAGCGCGGCCGAGTCC 48

seq_name: gb_pat:AR112421

seq_documentation_block:
LOCUS AR112421 237 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 11 from patent US 6130065.
ACCESSION AR112421
VERSION AR112421.1 GI:14092321

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 237)

AUTHORS Belt,J.A. and Crawford,C.R.
TITLE Nitrobenzylmercaptopyridine riboside (NEMPR)-insensitive,
equilibrative, nucleoside transport protein, nucleic acids encoding
the same and methods of use
JOURNAL Patent: US 6130065-A 11 10-OCT-2000;
FEATURES Location/Qualifiers
source 1..237

BASE COUNT 35 a 72 c 94 g 36 t
ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AR112421 ..

Align seg 1/1 to: AR112421 from: 1 to: 237

1 GluAlaGly*****Ser 6
|||||
193 GAGCGCGGTGCGTCTCG 210

seq_name: gb_pl:CNS01B2X

seq_documentation_block:

LOCUS CNS01B2X 240 bp mRNA linear PLN 02-SEP-1999
DEFINITION Botrytis cinerea strain T4 cDNA library under conditions of
nitrogen deprivation..

ACCESSION AL113713
VERSION AL113713.1 GI:5828332
KEYWORDS cDNA library; nitrogen deprivation.
SOURCE Botryotinia fuckeliana.
ORGANISM Botryotinia fuckeliana

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
Helotiales; Sclerotiniaceae; Botryotinia.

REFERENCE 1 (bases 1 to 240)

AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
78026 Versailles, France

REFERENCE 2 (bases 1 to 240)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
CP 5706 91057 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr
The cDNA library to be analyzed within the framework of this
project was created using a Botrytis cinerea strain which was grown
under conditions of nitrogen deprivation, which is the normal
situation for B. cinerea during its development on its host plant.

The library was produced in an oriented direction, in the pBSII
vector.

FEATURES
source

Location/Qualifiers
1..240
/organism="Botryotinia fuckeliana"
/strain="T4"
/db_xref="taxon:40559"
/note="Genoscope sequence ID : W53H081"

BASE COUNT 70 a 37 c 81 g 52 t
ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x CNS01B2X ..

Align seg 1/1 to: CNS01B2X from: 1 to: 240

1 GluAlaGly*****Ser 6

|||||
28 GAGCGGGGCCACGAGC 45

seq_name: gb_ba:AF068239

seq_documentation_block:

LOCUS AF068239 245 bp DNA linear BCT 14-JUN-1998
DEFINITION Bacterium Clab3 dichloro-catechol 1,2-dioxygenase (tfdc) gene,
partial cds.

ACCESSION AF068239

VERSION AF068239.1 GI:3219707

KEYWORDS

SOURCE bacterium Clab3.

ORGANISM Bacterium Clab3

REFERENCE 1 (bases 1 to 245)

AUTHORS Leander,M., Vallaeys,T. and Fulthorpe,R.R.
TITLE Primers for amplification and determination of the chloro- and
dichloro-catechol 1,2-dioxygenases family
JOURNAL Can. J. Microbiol. (1998) in press

REFERENCE 2 (bases 1 to 245)

AUTHORS Vallaeys,T., Courde,L., McGowan,C., Wright,A.D. and Fulthorpe,R.R.
TITLE The construction of catabolic pathway for the degradation of the
herbicide 2,4-D involves independent recruitment of gene cassettes

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 245)

AUTHORS Leander,M., Vallaeys,T., Fulthorpe,R.R. and Courde,L.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1998) Microbiologie des Sols, INRA, 17 rue Sully,
Dijon 21034, France

FEATURES
source

Location/Qualifiers
1..245
/organism="bacterium Clab3"
/strain="Clab3"
/db_xref="taxon:77202"
<1..>245
/gene="tfdc"

gene

CDS

<1..>245

/gene="tfdc"

/function="2,4-D pesticide degradation"

/codon_start=3

/transl_table=11

/product="dichloro-catechol 1,2-dioxygenase"

/protein_id="AAC23509.1"

/db_xref="GI:3219708"

/translation="HRTPDGKYSGFHDGIPDSFYGRGLSVGPDGGSFRVRRTPVAYQI
PNDGPTGALLETMGGHSHWRPAHFVKAPQVETLT"

BASE COUNT

ORIGIN

47 a 78 c 79 g 41 t

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AF068239 ..

Align seg 1/1 to: AF068239 from: 1 to: 245

1 GluAlaGly*****Ser 6

|||||

64 GAGCAGGCTCAGCGTCG 81

seq_name: gb_sts:G59419

seq_documentation_block:

LOCUS G59419 255 bp DNA linear STS 30-MAR-2000
 DEFINITION SHGC-130130 Human Homo sapiens STS genomic, sequence tagged site.

ACCESSION G59419

VERSION G59419.1 GI:6124738

KEYWORDS

STS.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 255)

Oliver,M. and Cox,D.R.

Unpublished, Olivier, M., Cox, D.R. (2000)

Unpublished

COMMENT

Contact: Michael Olivier, David R. Cox

Stanford Human Genome Center

Stanford University School of Medicine

4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA

Tel: (650) 320-5800

Fax: (650) 320-5801

Email: olivier@shgc.stanford.edu

Primer A: TCGAAGGCTGCTGTTTATC

Primer B: GCTTGTGATTCTCTGGAGGT

STS size: 207

PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds

Annealing: 60 degrees C for 30 seconds

Polymerization: 72 degrees C for 23 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Amplitaq Gold Polymerase: 0.07 units/ul

Total Vol: 5 ul

Buffer:

MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 8.3

Finished human sequences in NCBI. STSs designed and developed at the

Stanford Human Genome Center.

Location/Qualifiers

source

1. .255

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="12"

/clone_lib="Human"

32. .238

primer_bind

primer_bind complement(216) .238)

BASE COUNT 55 a 59 c 64 g 77 t
 ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x G59419 ..

Align seg 1/1 to: G59419 from: 1 to: 255

1 GluAlaGly*****Ser 6

|||||

12 GAGGCTGGGCAAGTCG 29

seq_name: gb_pr:HSU87908

seq_documentation_block:

LOCUS HSU87908 257 bp DNA linear PRI 31-JUL-1998
 DEFINITION Human aconitase hydratase (ACO2) gene, exon 9.

ACCESSION U87933

VERSION U87933.1

GI:3366612

KEYWORDS

8 of 14

SEGMENT

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 257)

Mirel,D.B., Marder,K., Graziano,J., Freyer,G., Zhao,Q., Mayeux,R.

and Wilhelmsen,K.C.

Characterization of the human mitochondrial aconitase gene

Gene 213 (1-2), 205-218 (1998)

98296268

REFERENCE

2 (bases 1 to 257)

Mirel,D.B. and Wilhelmsen,K.C.

Direct Submission

Submitted (30-JAN-1997)

Ernest Gallo Clinic and Research Center,

University of California, San Francisco, SFGH Bldg1, Room 101, 1001

Potrero Avenue, San Francisco, CA 94110, USA

Location/Qualifiers

1. .257

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="22"

/map="22q13"

<1. .69

/gene="ACO2"

/number=8

70. .175

/gene="ACO2"

/number=9

176. .>257

/gene="ACO2"

/number=9

BASE COUNT 48 a 79 c 65 g 65 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x HSU87908/rev ..

Align seg 1/1 to reverse of: HSU87908 from: 1 to: 257

1 GluAlaGly*****Ser 6

```
|||||
19 GAAGCGGGGACAGCCTCT 2

seq_name: gb_sts:G02071

seq_documentation_block:
LOCUS       G02071             262 bp    DNA    linear    STS 03-OCT-1995
DEFINITION  human STS 662AR, sequence tagged site.
ACCESSION   G02071
VERSION     G02071.1  GI:632474
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 262)
AUTHORS     Miller,A.P., Gustashaw,K., Wolff,D.J., Rider,S.H., Monaco,A.P.,
             Ebler,B., Schlessinger,D., Gorski,J.L., van Ommen,G., Weissenbach,J.
             and Willard,H.F.
TITLE       Three genes that escape X chromosome inactivation are clustered
             within a 6 Mb YAC contig and STS map in Xp11.21-p11.22
JOURNAL     Hum. Mol. Genet. 4 (4), 731-739 (1995)
MEDLINE     95359982
PUBMED      7633424
COMMENT     Contact: Andrew P. Miller
             Genetics
             Case Western Reserve University Willard Lab
             2109 Adelbert Road, Cleveland, OH 44106, USA
             Tel: 216-368-3518
             Fax: 216-368-3432
             Email: apm6@po.cwru.edu

Primer A: CTCGCTCTCTCTTCTCTCC
Primer B: TTTTCAAGGTGTGAAGG
STS size: 188
PCR profile:
  Presoak: 95 degrees C for 2.00 minute(s)
  Denaturation: 94 degrees C for 0.25 minute(s)
  Annealing: 55 degrees C for 0.25 minute(s)
  Polymerization: 72 degrees C for 0.25 minute(s)
  PCR Cycles: 30
  Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 20 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 25 ul

Buffer:
MgCl2: 2.0 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

Sequence is from the right end of YAC yHCEPH62g2.
FEATURES             Location/Qualifiers
     source            1..262
                       /organism="Homo sapiens"
                       /db_xref="taxon:9606"
                       /map="Xp11.21"
     STS               21..208
     primer_bind       21..41
     primer_bind       complement(189..208)
     BASE COUNT       61 a 84 c 52 g 65 t
     ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent similarity: 100.000 Percent identity: 100.000

alignment_block:
BASK-853-CLAIM4 x G02071/rev ..
Align seg 1/1 to reverse of: G02071 from: 1 to: 262

seq_name: gb_sts:G65760

seq_documentation_block:
LOCUS       G65760             262 bp    DNA    linear    STS 13-APR-2001
DEFINITION  Dazl Miscellaneous mouse cDNA sequences Mus musculus STS genomic,
             sequence tagged site.
ACCESSION   G65760
VERSION     G65760.1  GI:13624250
KEYWORDS    STS.
SOURCE      house mouse.
ORGANISM    Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 262)
AUTHORS     Wang,P.J., McCarrey,J.R., Yang,F. and Page,D.C.
TITLE       An abundance of X-linked genes expressed in spermatogonia
JOURNAL     Nat. Genet. 27 (4), 422-426 (2001)
MEDLINE     21175748
PUBMED      11279525
COMMENT     Synonyms: JW109/JW110
             Contact: Jeremy Wang
             Page Lab
             Whitehead Institute for Biomedical Research
             9 Cambridge Center, Cambridge, MA 02142, USA
             Tel: 6172588420
             Fax: 6172586321
             Email: pjwang@wi.mit.edu
             Primer A: TTCAGGCATATCCTCCTTATC
             Primer B: ATGCTTCGTCACAGACTTC
             STS size: 262
             PCR profile:
               Presoak: 94C for 2 minutes
               Denaturation: 94C for 30 seconds
               Annealing: 56C for 30 seconds
               Polymerization: 72C for 2 minutes
               PCR Cycles: 30 or 35
               Thermal Cycler: Hybaid Omn-E

Protocol:
Template: 2 or 5 ng total oligo(dT)18 or random primed cDNAs
Primer: each 1 uM
dNTPs: 100 uM
Taq Polymerase: 1 unit
Total Vol: 25 ul

Buffer:
MgCl2: 1.5 mM
NH4Cl: 5.0 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.2 at 25oC

RT-PCR assay.
FEATURES             Location/Qualifiers
     source            1..262
                       /organism="Mus musculus"
                       /db_xref="taxon:10090"
                       /clone_lib="Miscellaneous mouse cDNA sequences"
                       /note="Mouse STSs derived from miscellaneous cDNA
                       sequences"
     STS               1..262
     primer_bind       1..21
     primer_bind       complement(242..262)
     BASE COUNT       67 a 64 c 59 g 72 t
```

ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x G65760/rev ..

Align seg 1/1 to reverse of: G65760 from: 1 to: 262

1 GluAlaGly*****Ser 6

|||||

223 GAGCTGGAGCAGCATCA 206

seq_name: gb_pr:HS52G1R

seq_documentation_block:

LOCUS HS52G1R 263 bp DNA linear PRI 22-OCT-1995
 DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 52g1, reverse
 read cp952g1.rta.

ACCESSION 261462

VERSION 261462.1 GI:1033840

KEYWORDS CpG island; genomic MseI fragment.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 263)

MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.

Direct Submission

Submitted (16-OCT-1995). The Sanger Centre, Hinxton, Cambridgeshire,

CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk

2 (bases 1 to 263)

Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.

Purification of CpG islands using a methylated DNA binding column

Nat. Genet. 6 (3), 236-244 (1994)

MEDLINE 94282070

COMMENT

Vector: pGEM-5zf(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
<http://www.hgmp.mrc.ac.uk/> for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES

source

1..263
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /sex="male"
 /dev_stage="adult"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /clone="52g1"

BASE COUNT 49 a 88 c 84 g 41 t 1 others
 ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x HS52G1R ..

Align seg 1/1 to: HS52G1R from: 1 to: 263

1 GluAlaGly*****Ser 6

|||||

54 GAGCGGGGGCTGCATCC 71

seq_name: gb_om:HLU03593

seq_documentation_block:

LOCUS HLU03593 264 bp DNA linear MAM 18-JAN-1995
 DEFINITION Hydrurga leptonyx histone H2AF processed pseudogene, partial
 sequence.

ACCESSION U03593

VERSION U03593.1 GI:458188

KEYWORDS

SOURCE leopard seal.

ORGANISM

Hydrurga leptonyx
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Carnivora; Pinnipedia; Phocidae; Hydrurga.

REFERENCE

1 (bases 1 to 264)

Slade,R.W., Moritz,C. and Heideman,A.

Multiple nuclear-gene phylogenies: application to pinnipeds and

comparison with a mitochondrial DNA gene phylogeny

Mol. Biol. Evol. 11 (3), 341-356 (1994)

MEDLINE 94285805

REFERENCE 2 (bases 1 to 264)

Slade,R.W.

Direct Submission

Submitted (18-NOV-1993) Robert W. Slade, Immunogenetik,

Max-Planck-Institut fuer Biologie, Corrensstrasse 42, D-72076

TUEBINGEN, Germany

FEATURES

source

1..264

/organism="Hydrurga leptonyx"

/db_xref="taxon:29086"

<1..>264

/note="histone H2AF processed pseudogene"

BASE COUNT 72 a 71 c 68 g 53 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x HLU03593/rev ..

Align seg 1/1 to reverse of: HLU03593 from: 1 to: 264

1 GluAlaGly*****Ser 6

|||||

226 GAGCTGGTTCTCTCT 209

seq_name: gb_om:HLU03595

seq_documentation_block:

LOCUS HLU03595 264 bp DNA linear MAM 18-JAN-1995
 DEFINITION Leptonychotes weddelli histone H2AF processed pseudogene, partial
 sequence.

ACCESSION U03595

VERSION U03595.1 GI:458191

KEYWORDS

SOURCE Weddell seal.

ORGANISM

Leptonychotes weddelli
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Carnivora; Pinnipedia; Phocidae; Leptonychotes.

REFERENCE

1 (bases 1 to 264)

Slade,R.W., Moritz,C. and Heideman,A.

Multiple nuclear-gene phylogenies: application to pinnipeds and

comparison with a mitochondrial DNA gene phylogeny

Mol. Biol. Evol. 11 (3), 341-356 (1994)

MEDLINE 94285805

REFERENCE 2 (bases 1 to 264)

Slade,R.W.

Direct Submission

Submitted (18-NOV-1993) Robert W. Slade, Immunogenetik,

Max-Planck-Institut fuer Biologie, Corrensstrasse 42, D-72076

TUEBINGEN, Germany

```

FEATURES
  source      Location/Qualifiers
    1..264
    /organism="Leptonychotes weddellii"
    /db_xref="taxon:9713"
  misc_feature <1..>264
    /note="histone H2AF processed pseudogene"
  BASE COUNT  72 a  71 c  68 g  53 t
  ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x LMU03595/rev ..
  Align seg 1/1 to reverse of: LMU03595 from: 1 to: 264
    1 GluAlaGly*****Ser 6
    |||||
  226 GAGCGTGGTCTTCCTCT 209

seq_name: gb_om:MAU03600

seq_documentation_block:
  LOCUS MAU03600 264 bp DNA linear MAM 18-JAN-1995
  DEFINITION Mirounga angustirostris histone H2AF processed pseudogene, partial
  sequence.
  ACCESSION U03600
  VERSION U03600.1 GI:458199
  KEYWORDS
  SOURCE northern elephant seal.
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
  REFERENCE
    1 (bases 1 to 264)
    Slade,R.W., Moritz,C. and Heideman,A.
    Multiple nuclear-gene phylogenies: application to pinnipeds and
    comparison with a mitochondrial DNA gene phylogeny
  TITLE
    Mol. Biol. Evol. 11 (3), 341-356 (1994)
  JOURNAL
  MEDLINE 94285805
  REFERENCE 2 (bases 1 to 264)
  AUTHORS Slade,R.W.
  TITLE Direct Submission
  JOURNAL Submitted (18-NOV-1993) Robert W. Slade, Immunogenetik,
  Max-Planck-Institut fuer Biologie, Corrensstrasse 42, D-72076
  Tuebingen, Germany

FEATURES
  source      Location/Qualifiers
    1..264
    /organism="Phoca vitulina"
    /db_xref="taxon:9720"
  misc_feature <1..>264
    /note="histone H2AF processed pseudogene"
  BASE COUNT  74 a  69 c  68 g  53 t
  ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x PVU03589/rev ..
  Align seg 1/1 to reverse of: PVU03589 from: 1 to: 264
    1 GluAlaGly*****Ser 6
    |||||
  226 GAGCGTGGTCTTCCTCT 209

seq_name: gb_pr:HS121G1F

seq_documentation_block:
  LOCUS HS121G1F 265 bp DNA linear PRI 17-OCT-1995
  DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 121g1,
  forward read cpgl21g1.ftla.
  ACCESSION Z54496
  VERSION Z54496.1 GI:1020537
  KEYWORDS CpG island; genomic MseI fragment.
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  REFERENCE 1 (bases 1 to 265)
  AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Mickle,G.
  TITLE Direct Submission
  JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
  REFERENCE 2 (bases 1 to 265)
  AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE Purification of CpG islands using a methylated DNA binding column
  JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
  MEDLINE 94282070
  COMMENT
    Vector: pGEM-5zf(-)
    Clones are available from the UK MRC Human Genome Mapping Project

```

```

seq_name: gb_om:PVU03589

seq_documentation_block:
  LOCUS PVU03589 264 bp DNA linear MAM 18-JAN-1995
  DEFINITION Phoca vitulina histone H2AF processed pseudogene, partial sequence.
  ACCESSION U03589
  VERSION U03589.1 GI:458183
  KEYWORDS
  SOURCE harbor seal.
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Phoca.
  REFERENCE 1 (bases 1 to 264)
    Slade,R.W., Moritz,C. and Heideman,A.
    Multiple nuclear-gene phylogenies: application to pinnipeds and
    comparison with a mitochondrial DNA gene phylogeny
  TITLE
    Mol. Biol. Evol. 11 (3), 341-356 (1994)
  JOURNAL
  MEDLINE 94285805
  REFERENCE 2 (bases 1 to 264)
  AUTHORS Slade,R.W.
  TITLE Direct Submission
  JOURNAL Submitted (18-NOV-1993) Robert W. Slade, Immunogenetik,
  Max-Planck-Institut fuer Biologie, Corrensstrasse 42, D-72076
  Tuebingen, Germany

FEATURES
  source      Location/Qualifiers
    1..264
    /organism="Phoca vitulina"
    /db_xref="taxon:9720"
  misc_feature <1..>264
    /note="histone H2AF processed pseudogene"
  BASE COUNT  74 a  69 c  68 g  53 t
  ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x PVU03589/rev ..
  Align seg 1/1 to reverse of: PVU03589 from: 1 to: 264
    1 GluAlaGly*****Ser 6
    |||||
  226 GAGCGTGGTCTTCCTCT 209

seq_name: gb_pr:HS121G1F

seq_documentation_block:
  LOCUS HS121G1F 265 bp DNA linear PRI 17-OCT-1995
  DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 121g1,
  forward read cpgl21g1.ftla.
  ACCESSION Z54496
  VERSION Z54496.1 GI:1020537
  KEYWORDS CpG island; genomic MseI fragment.
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  REFERENCE 1 (bases 1 to 265)
  AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Mickle,G.
  TITLE Direct Submission
  JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
  REFERENCE 2 (bases 1 to 265)
  AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE Purification of CpG islands using a methylated DNA binding column
  JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
  MEDLINE 94282070
  COMMENT
    Vector: pGEM-5zf(-)
    Clones are available from the UK MRC Human Genome Mapping Project

```

```

Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
  source
    1..265
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /sex="male"
      /dev_stage="adult"
      /tissue_type="blood"
      /clone_lib="CGI-1"
      /clone="121g1"
BASE COUNT      47 a   83 c   88 g   41 t   6 others
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x H5121G1F/rev ..
Align seg 1/1 to reverse of: H5121G1F/rev from: 1 to: 265

1 GluAlaGly*****Ser 6
|||||
157 GAGCGCGGGCGGCGCTCT 140

seq_name: gb_ro:MUSIGFBP02

seq_documentation_block:
LOCUS      MUSIGFBP02               268 bp      DNA      linear      ROD 05-SEP-1995
DEFINITION Mouse insulin-like growth factor binding protein 2 (IGFBP-2) gene,
            exon 2.
ACCESSION  L05437
VERSION    L05437
KEYWORDS   L05437
SEGMENT    2 of 4
SOURCE     mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 268)
AUTHORS    Landwehr,J., Kaupmann,K., Heinrich,G. and Schwander,J.
TITLE      Cloning and characterization of the gene encoding murine
            insulin-like growth factor-binding protein-2, mIGFBP-2
JOURNAL    Gene 124 (2), 281-286 (1993)
MEDLINE    93185936
COMMENT     On Sep 5, 1995 this sequence version replaced gi:194491.
FEATURES
  source
    1..268
      /organism="Mus musculus"
      /strain="BALB/c"
      /sub_species="domesticus"
      /db_xref="taxon:10090"
      /map="proximal region of chromosome 1"
      /sex="female"
      /tissue_type="fetal liver"
    23..249
      /gene="IGFBP-2"
      /number=2
      /evidence="experimental"
BASE COUNT      62 a   66 c   94 g   46 t
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM4 x MUSIGFBP02/rev ..
Align seg 1/1 to reverse of: MUSIGFBP02 from: 1 to: 268

1 GluAlaGly*****Ser 6
|||||
241 GAGCGCGGGCGCAACTTCT 224

seq_name: gb_pr:HS121G1R

seq_documentation_block:
LOCUS      HS121G1R               269 bp      DNA      linear      PRI 17-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 121g1,
            reverse read cp9121g1.rt1a.
ACCESSION  Z54497
VERSION    Z54497.1 GI:1020538
KEYWORDS   CpG island; genomic MseI fragment.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 269)
AUTHORS    MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE      Direct Submission
JOURNAL    Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
REFERENCE  2 (bases 1 to 269)
AUTHORS    Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE      Purification of CpG islands using a methylated DNA binding column
            Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE    94282070
COMMENT     Vector: pGEM-5zf(-)
            Clones are available from the UK MRC Human Genome Mapping Project
            Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
            http://www.hgmp.mrc.ac.uk/ for details
            or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
  source
    1..269
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /sex="male"
      /dev_stage="adult"
      /tissue_type="blood"
      /clone_lib="CGI-1"
      /clone="121g1"
BASE COUNT      43 a   89 c   88 g   47 t   2 others
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x HS121G1R ..
Align seg 1/1 to: HS121G1R from: 1 to: 269

1 GluAlaGly*****Ser 6
|||||
110 GAGCGCGGGCGGCGCTCT 127

seq_name: gb_pr:HSACO2G08

seq_documentation_block:
LOCUS      HSACO2G08               270 bp      DNA      linear      PRI 11-MAR-1999
DEFINITION Homo sapiens aconitase (ACO2) gene, nuclear gene encoding
            mitochondrial protein, exon 9.
ACCESSION  AF093087
VERSION    AF093087.1 GI:4378036
KEYWORDS

```

```

SEGMENT      8 of 17
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 270)
JOURNAL      Huang,H.H.
REFERENCE    Characteristic and location of human mitochondrial aconitase gene
AUTHORS      Unpublished
TITLE        2 (bases 1 to 270)
JOURNAL      Huang,H.H.
TITLE        Direct Submission
JOURNAL      Submitted (16-SEP-1998) Anatomy, Chang Gung University, 259 Wen-Hua
FEATURES     1st Road, Kwei-Shan, Tao-Yuan, Taiwan
SOURCE       1. .270
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="22"
              /map="22q13.22"
              /sex="male"
              /cell_type="circulating lymphocytes"
              83. .188
              /gene="ACO2"
              /number=9
BASE COUNT   51 a 80 c 72 g 67 t
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x HSACO2G08/rev ..
Align seg 1/1 to reverse of: HSACO2G08 from: 1 to: 270

1 GluAlaGly*****Ser 6
|||||
32 GAAGCGGGGACAGCCTCT 15

seq_name: gb_pat:162369

seq_documentation_block:
LOCUS      162369      272 bp      DNA      linear      PAT 07-OCT-1997
DEFINITION Sequence 10 from patent US 5658882.
ACCESSION  162369
VERSION    162369.1 GI:2480317
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 272)
AUTHORS    Celeste,A.J., Wozney,J.M., Rosen,V.A., Wolfman,N.M., Thomsen,G.H.
           and Melton,D.A.
TITLE      Methods of inducing formation of tendon and/or ligament tissue
           comprising administering BMP-12, BMP-13, and/or MP-52
JOURNAL    Patent: US 5658882-A 10 19-AUG-1997;
FEATURES   Location/Qualifiers
SOURCE     1. .272
           /organism="unknown"
BASE COUNT 57 a 91 c 72 g 52 t
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

```

```

BASK-853-CLAIM4 x I62369/rev ..
Align seg 1/1 to reverse of: I62369 from: 1 to: 272

1 GluAlaGly*****Ser 6
|||||
170 GAGGCTGGCGCAGCGTCG 153

seq_name: gb_om:MIRHIS2APP

seq_documentation_block:
LOCUS      MIRHIS2APP      274 bp      DNA      linear      MAM 08-JUL-1994
DEFINITION Mirounga leonina histone H2AF processed pseudogene.
ACCESSION  LI7502
VERSION    LI7502.1 GI:308965
KEYWORDS   histone H2AF; pseudogene.
SOURCE     Mirounga leonina DNA.
ORGANISM   Mirounga leonina
           Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
REFERENCE  1 (bases 1 to 274)
AUTHORS    Slade,R.W., Moritz,C., Heideman,A. and Hale,P.T.
TITLE      Rapid assessment of single-copy nuclear DNA variation in diverse
           species
JOURNAL    Mol. Ecol. 2, 359-373 (1993)
MEDLINE    94214718
FEATURES   Location/Qualifiers
SOURCE     1. .274
           /organism="Mirounga leonina"
           /db_xref="taxon:9715"
           <1..>274
           /codon_start=1
           /pseudo
BASE COUNT 75 a 75 c 71 g 53 t
ORIGIN

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x MIRHIS2APP/rev ..
Align seg 1/1 to reverse of: MIRHIS2APP from: 1 to: 274

1 GluAlaGly*****Ser 6
|||||
228 GAGGCTGGTTCTCTCTCT 211

seq_name: gb_sts:G64229

seq_documentation_block:
LOCUS      G64229      275 bp      DNA      linear      STS 24-OCT-2000
DEFINITION P288H2/T7 Human Chromosome 12 Homo sapiens STS genomic clone
ACCESSION  RPC11-288H2 T7, sequence tagged site.
VERSION    G64229
KEYWORDS   G64229.1 GI:9802441
SOURCE     STS.
           human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 275)
AUTHORS    Lee,H., Choi,E., Seomun,Y., Montgomery,K., Huebner,A., Lee,E.,
           Lau,S., Joo,C.K., Kucherlapati,R. and Yoon,S.J.
TITLE      High-resolution transcript map of the region spanning D12S1629 and
           D12S312 at chromosome 12q13: triple A syndrome-linked region
JOURNAL    Genome Res. 10 (10), 1561-1567 (2000)
MEDLINE    20499368
COMMENT

```

Contact: Sung-Joo Kim Yoon

Sung-Joo Kim Lab., Research Institute of Molecular Genetics
Catholic Research Institutes of Medical Science
505 Banpo-dong, Seocho-Ku, Seoul 137-040 Korea
Tel: 0118225902604
Fax: 0118225902603
Email: sjkyoon@cmc.cuk.ac.kr
Primer A: TCTATGGGGAACGTCCTAACCC
Primer B: CAGCACCTACTTCTGCCCA
STS size: 206

Protocol:
template 30-100ng
each primer each 1mM
dNTP 200 mM
TAO Gold 0.75U
MgCl2 1.5mM
BSA 30 mg/ml
Final Volume 20
Preheat: 12 minutes 95 degrees C
Denaturation: 25 seconds 94 degrees C
Annealing: 25 seconds 57 degrees C
Polymerization: 10 seconds 72 degrees C
PCR Cycles: 36

Buffer:
KCl: 50 mM
Tris-HCl: 10 mM
PH: 8.3

PAC 288H2 from RPCI-1 human PAC Library
(http://bacpac.med.buffalo.edu/), maps between AFMa283yg1 and
AFMa207va9, 12q13.

FEATURES
source Location/Qualifiers
1..275
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RPC11-288H2"
/clone_lib="Human Chromosome 12"
43..248
STS
primer_bind 43..63
primer_bind complement(230..248)
BASE COUNT 46 a 83 c 90 g 56 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x G64229/rev ..
Align seg 1/1 to reverse of: G64229 from: 1 to: 275

1 GluAlaGly*****Ser 6
|||||
168 GAGCGAGGGCGAGCCAGC 151

seq_name: gb_pat:AR084752

seq_documentation_block:
LOCUS AR084752 276 bp DNA linear PAT 01-SEP-2000
DEFINITION Sequence 78 from patent US 5981218.
ACCESSION AR084752
VERSION AR084752.1 GI:10011523
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE
1 (bases 1 to 276)
AUTHORS Rio,M.-C., Tomasetto,C., Basset,P. and Byrne,J.
TITLE Isolated nucleic acid molecules useful as leukemia markers and in
breast cancer prognosis and encoded polypeptides

JOURNAL Patent: US 5981218-A 78 09-NOV-1999;
FEATURES Location/Qualifiers
source 1..276
/organism="unknown"
BASE COUNT 94 a 60 c 53 g 61 t 8 others
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x AR084752/rev ..
Align seg 1/1 to reverse of: AR084752 from: 1 to: 276

1 GluAlaGly*****Ser 6
|||||
184 GAGCGAGGGCGAGTCATCT 167

THIS PAGE BLANK (USPTO)

OM of: BASK-853-CLAIM4 to: Issued_Patents_NA:* out_format : pfs

Date: Sep 24, 2002 2:13 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/BASKR853079/runat_17092002_103408_1838/app_query.fasta_1.98
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=mod.rni
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
-LOOPEXT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000
-XGAPEXT=0.500 -XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000
-MATRIX=blossum62 -TRANS=human40.cdi -LIST=50 -DOCALLIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=BASKR853079_@CGN1_1.72 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: BASK-853-CLAIM4

Query length: 6

Database: Issued_Patents_NA:*

Database sequences: 383533

Database length: 122816752

Search time (sec): 82.560000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-887-798-19	-	19.00	113.19	128.80	26
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-849-602C-28	-	19.00	111.59	158.07	32
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-244-378A-20	-	19.00	111.13	167.81	34
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-244-378A-58	-	19.00	111.13	167.81	34
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-151-574-41	+	19.00	110.69	177.54	36
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-151-574-42	+	19.00	110.69	177.54	36
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-419-448-41	+	19.00	110.69	177.54	36
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-419-448-42	+	19.00	110.69	177.54	36
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-233-510-41	+	19.00	110.69	177.54	36
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-233-510-42	+	19.00	110.69	177.54	36
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-374-652C-94	+	19.00	104.17	409.39	84
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-058-389A-11	+	19.00	96.20	1.1e+03	237
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-362-670B-10	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-333-576C-10	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-10	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-11	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-12	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-13	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-14	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-15	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-16	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-17	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-18	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-19	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-20	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-21	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-22	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-23	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-24	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-25	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-26	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-27	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-28	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-29	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-30	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-31	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-32	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-33	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-34	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-35	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-36	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-37	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-38	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-39	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-40	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-41	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-42	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-43	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-44	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-45	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-46	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-47	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-48	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-49	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-50	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-51	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-52	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-53	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-54	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-55	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-56	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-57	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-58	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-59	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-60	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-61	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-62	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-63	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-64	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-65	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-66	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-67	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-68	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-69	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-70	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-71	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-72	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-73	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-74	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-75	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-76	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-77	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-78	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-79	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-80	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-81	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-82	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-83	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-84	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-85	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-86	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-87	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-88	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-89	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-90	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-91	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-92	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-93	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-94	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-95	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-96	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-97	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-98	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-99	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-100	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-101	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-102	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-103	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-104	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-105	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-106	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-107	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-108	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-109	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-110	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-111	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-112	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-113	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-114	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-115	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-116	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-117	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-118	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-119	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-120	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-121	-	19.00	95.14	1.3e+03	272
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-122	-	19.00	95		

```
seq_documentation_block:
; Sequence 28, Application US/08849602C
; Patent No. 6171823
; GENERAL INFORMATION:
; APPLICANT: Fabricius, Helle
; APPLICANT: Hastrup, Sven
; TITLE OF INVENTION: A Process of Producing
; TITLE OF INVENTION: Extracellular Proteins in Bacteria
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6171823o No. 6171823disk of No. 6171823th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,602C
; FILING DATE: 02-JUN-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4139.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-849-602C-28

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-849-602C-28/rev ..
Align seg 1/1 to reverse of: US-08-849-602C-28 from: 1 to: 32

1 GluAlaGly*****Ser 6
|||||
30 GAGCGCGGCGCGGAGC 13

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-244-378A-20

seq_documentation_block:
; Sequence 20, Application US/08244378A
; Patent No. 5756323
; GENERAL INFORMATION:
; APPLICANT: Kallenbach, Sacha
; APPLICANT: Doyen, No. 5756323lle
; APPLICANT: Rougeon, Francois
; TITLE OF INVENTION: METHOD FOR GENERATING STRUCTURAL AND
; TITLE OF INVENTION: FUNCTIONAL DIVERSITY IN A PEPTIDE SEQUENCE
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
```

```
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,378A
; FILING DATE: 01-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91/15389
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5756323man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-078-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-244-378A-20

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-244-378A-20/rev ..
Align seg 1/1 to reverse of: US-08-244-378A-20 from: 1 to: 34

1 GluAlaGly*****Ser 6
|||||
28 GAGCGCGGATCCACTAGT 11

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-244-378A-58

seq_documentation_block:
; Sequence 58, Application US/08244378A
; Patent No. 5756323
; GENERAL INFORMATION:
; APPLICANT: Kallenbach, Sacha
; APPLICANT: Doyen, No. 5756323lle
; APPLICANT: Rougeon, Francois
; TITLE OF INVENTION: METHOD FOR GENERATING STRUCTURAL AND
; TITLE OF INVENTION: FUNCTIONAL DIVERSITY IN A PEPTIDE SEQUENCE
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,378A
; FILING DATE: 01-SEP-1994
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91/15389
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5756323man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-078-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-244-378A-58

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-244-378A-58/rev ..
Align seg 1/1 to reverse of: US-08-244-378A-58 from: 1 to: 34

1 GluAlaGly*****Ser 6
|||||
28 GAGCGCGGATCCACTAGT 11

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-151-574-41

seq_documentation_block:
; Sequence 41, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
```

```
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE: 18-2
; ORIGINAL SOURCE: US-08-151-574-41

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-151-574-41 ..
Align seg 1/1 to: US-08-151-574-41 from: 1 to: 36

1 GluAlaGly*****Ser 6
|||||
2 GAGCGCGGACTGCCAGT 19

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-151-574-42

seq_documentation_block:
; Sequence 42, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 42:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; US-08-151-574-42
;
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM4 x US-08-151-574-42/rev ..
Align seg 1/1 to reverse of: US-08-151-574-42 from: 1 to: 36
1 GluAlaGly*****Ser 6
|||||
35 GAGCGGGGACTGCCAGT 18
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-419-448-41
seq_documentation_block:
; Sequence 41, Application US/08419448
; Patent No. 5863533
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
```

```
; ORIGINAL SOURCE: 18-2
; US-08-419-448-41
;
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM4 x US-08-419-448-41 ..
Align seg 1/1 to: US-08-419-448-41 from: 1 to: 36
1 GluAlaGly*****Ser 6
|||||
2 GAGCGGGGACTGCCAGT 19
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-419-448-42
seq_documentation_block:
; Sequence 42, Application US/08419448
; Patent No. 5863533
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; US-08-419-448-42
;
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:
BASK-853-CLAIM4 x US-08-419-448-42/rev ..
Align seg 1/1 to reverse of: US-08-419-448-42 from: 1 to: 36

1 GluAlaGly*****Ser 6
|||||
35 GAGCGGGGACTGCCAGT 18
seq_name: /cgn2_6/ptodata/2/ina/6B/COMB.seq:US-09-233-510-41

seq_documentation_block:
; Sequence 41, Application US/09233510
; Patent No. 6350602
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/233,510
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24515-20026.00
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE: 18-2
; US-09-233-510-41

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM4 x US-09-233-510-41 ..

Align seg 1/1 to: US-09-233-510-41 from: 1 to: 36
1 GluAlaGly*****Ser 6
|||||
2 GAGCGGGGACTGCCAGT 19

seq_name: /cgn2_6/ptodata/2/ina/6B/COMB.seq:US-09-233-510-42
seq_documentation_block:
; Sequence 42, Application US/09233510
; Patent No. 6350602
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/233,510
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE: 18-3
; US-09-233-510-42

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM4 x US-09-233-510-42/rev ..
Align seg 1/1 to reverse of: US-09-233-510-42 from: 1 to: 36
1 GluAlaGly*****Ser 6
|||||

```
35 GAGCGGGGACTGCCAGT 18
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-374-652C-94
seq_documentation_block:
; Sequence 94, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTATE DEGRADING
; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058
; FILING DATE: 27-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925,401
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1050.071001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-374-652C-94

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-374-652C-94
Align seg 1/1 to: US-08-374-652C-94 from: 1 to: 84
1 GluAlaGly*****Ser 6
```

```
|||||
3 GAGCGGGGACTGCCAGT 20
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-058-389A-11
seq_documentation_block:
; Sequence 11, Application US/09058389A
; Patent No. 6130065
; GENERAL INFORMATION:
; APPLICANT: Belt, Judith A.
; APPLICANT: Crawford, Charles R.
; APPLICANT: Patel, Divyen
; TITLE OF INVENTION: A NITROBENZYL MERCAPTO PURINERIBOSIDE
; TITLE OF INVENTION: (NEMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
; TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
; TITLE OF INVENTION: USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,389A
; FILING DATE: April 9, 1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-013N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "intron 1"
; HYPOTHETICAL: NO
US-09-058-389A-11

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-09-058-389A-11
Align seg 1/1 to: US-09-058-389A-11 from: 1 to: 237
1 GluAlaGly*****Ser 6
|||||
193 GAGCGGGTCCGTCCTCG 210

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-362-670B-10
seq_documentation_block:
; Sequence 10, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
```



```
;
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; APPLICANT: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: mouse
; IMMEDIATE SOURCE:
; CLONE: mvl
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..243
; US-08-362-670B-10

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-362-670B-10/rev ..
Align seg 1/1 to reverse of: US-08-362-670B-10 from: 1 to: 272

1 GluAlaGly*****Ser 6
|||||
170 GAGGCTGGCGCAGCGTCG 153

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-333-576C-10

seq_documentation_block:
; Sequence 10, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
```

```
;
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; APPLICANT: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,576C
; FILING DATE: No. 6027919ember 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: mouse
; IMMEDIATE SOURCE:
; CLONE: mvl
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..243
; US-08-333-576C-10

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-333-576C-10/rev ..
Align seg 1/1 to reverse of: US-08-333-576C-10 from: 1 to: 272

1 GluAlaGly*****Ser 6
|||||
170 GAGGCTGGCGCAGCGTCG 153

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-10

seq_documentation_block:
; Sequence 10, Application US/08808324
; Patent No. 6284872
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; APPLICANT: GENETICS INSTITUTE, INC.
```

```
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,324
; FILING DATE: Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: mouse
; IMMEDIATE SOURCE:
; CLONE: mv1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..243
; US-08-808-324-10
```

```
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
BASK-853-CLAIM4 x US-08-808-324-10/rev ..
```

```
Align seg 1/1 to reverse of: US-08-808-324-10 from: 1 to: 272
```

```
1 GluAlaGly*****Ser 6
|||||
170 GAGGCTGGCGCAGCGTCG 153
```

```
seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-14030A-10
```

```
seq_documentation_block:
; Sequence 10, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: mouse
; IMMEDIATE SOURCE:
; CLONE: mv1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..243
; PCT-US94-14030A-10
```

```
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
BASK-853-CLAIM4 x PCT-US94-14030A-10/rev ..
```

```
Align seg 1/1 to reverse of: PCT-US94-14030A-10 from: 1 to: 272
```

```
1 GluAlaGly*****Ser 6
|||||
170 GAGGCTGGCGCAGCGTCG 153
```

```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-691-814B-78
```

```
seq_documentation_block:
; Sequence 78, Application US/08691814B
; Patent No. 5981218
; GENERAL INFORMATION:
; APPLICANT: Rio, Marie-Christine
; APPLICANT: Tomasetto, Catherine
; APPLICANT: Basset, Paul
; APPLICANT: Byrne, Jennifer
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave, NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,814B
; FILING DATE: 31-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,183
; FILING DATE: 09-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1383.0090001
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2543
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; US-08-691-814B-78

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-691-814B-78/rev ..
Align seg 1/1 to reverse of: US-08-691-814B-78 from: 1 to: 276

1 GluAlaGly*****Ser 6
184 GAAGCGGCGAGTCATCT 167

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-540-242A-5

seq_documentation_block:
; Sequence 5, Application US/08540242A
; Patent No. 5827734
; GENERAL INFORMATION:
; APPLICANT: Weigle, David S.
; APPLICANT: Kuijper, Joseph L.
; APPLICANT: Forstrom, John W.
; APPLICANT: Lehner, Joyce M.
; TITLE OF INVENTION: APPETITE SUPPRESSION FACTOR AND RELATED METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,242A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-01C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
```

```
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..435
; US-08-540-242A-5

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-540-242A-5/rev ..
Align seg 1/1 to reverse of: US-08-540-242A-5 from: 1 to: 438

1 GluAlaGly*****Ser 6
295 GAGCGGGGAGCAGCTCT 278

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-01471-5

seq_documentation_block:
; Sequence 5, Application PC/TUS9601471
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: US
; APPLICANT: 98195
; TITLE OF INVENTION: APPETITE SUPPRESSION FACTOR AND RELATED METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01471
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A.
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-01PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..435
PCT-US96-01471-5

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x PCT-US96-01471-5/rev ..
  Align seg 1/1 to reverse of: PCT-US96-01471-5 from: 1 to: 438

      1 GluAlaGly*****Ser 6
      |||||
      295 GAGCAGGAGCAGCTCT 278

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-856-253-1

seq_documentation_block:
; Sequence 1, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: linear
US-08-856-253-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x PCT-US96-01471-5/rev ..
  Align seg 1/1 to reverse of: PCT-US96-01471-5 from: 1 to: 445

      1 GluAlaGly*****Ser 6
      |||||
      303 GAGCAGGAGCAGCTCT 286

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-540-242A-1

seq_documentation_block:
; Sequence 1, Application US/08540242A
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..435
PCT-US96-01471-5

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-856-253-1 ..
  Align seg 1/1 to: US-08-856-253-1 from: 1 to: 441

      1 GluAlaGly*****Ser 6
      |||||
      46 GAGCGGGAGACAGTAGT 63

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-688-908-6

seq_documentation_block:
; Sequence 6, Application US/08688908
; Patent No. 6287027
; GENERAL INFORMATION:
; APPLICANT: Spurlock, Michael E.
; TITLE OF INVENTION: BOVINE LEPTIN PROTEIN, NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHYTE HIRSCHBOECK DUDEK S.C.
; STREET: Sulte 2100 111 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,908
; FILING DATE: 31-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plotcher, Gary R.
; REGISTRATION NUMBER: 27,830
; REFERENCE/DOCKET NUMBER: PM-8808
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 414-273-2100
; TELEFAX: 414-223-5000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 445 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-688-908-6

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-688-908-6/rev ..
  Align seg 1/1 to reverse of: US-08-688-908-6 from: 1 to: 445

      1 GluAlaGly*****Ser 6
      |||||
      303 GAGCAGGAGCAGCTCT 286

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-540-242A-1

seq_documentation_block:
; Sequence 1, Application US/08540242A
```

```
; Patent No. 5827734
; GENERAL INFORMATION:
; APPLICANT: Weigle, David S.
; APPLICANT: Kuijper, Joseph L.
; APPLICANT: Forstrom, John W.
; APPLICANT: Lehner, Joyce M.
; TITLE OF INVENTION: APPETITE SUPPRESSION FACTOR AND RELATED METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,242A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-01C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..501
; US-08-540-242A-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-540-242A-1/rev ..
  Align seg 1/1 to reverse of: US-08-540-242A-1 from: 1 to: 504

1 GluAlaGly*****Ser 6
|||||
361 GAGGAGGAGGAGCTCT 344

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-692-922-6

seq_documentation_block:
; Sequence 6, Application US/08692922
; Patent No. 6277592
; GENERAL INFORMATION:
; APPLICANT: Bidwell, Christopher A.
; APPLICANT: Spurlock, Michael E.
; TITLE OF INVENTION: PORCINE LEPTIN PROTEIN, NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITE HIRSCHBOECK DUDEK S.C.
; STREET: Suite 2100 111 East Wisconsin Avenue
; CITY: Milwaukee
```

```
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/692,922
; FILING DATE: 31-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shaw, Melissa A.
; REGISTRATION NUMBER: 38,301
; REFERENCE/DOCKET NUMBER: PM-8935
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 414-273-2100
; TELEFAX: 414-223-5000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-692-922-6

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-692-922-6/rev ..
  Align seg 1/1 to reverse of: US-08-692-922-6 from: 1 to: 504

1 GluAlaGly*****Ser 6
|||||
361 GAGGAGGAGGAGCTCT 344

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-01471-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9601471
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: US
; APPLICANT: 98195
; TITLE OF INVENTION: APPETITE SUPPRESSION FACTOR AND RELATED METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: PCT/US96/01471
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawisak, Deborah A.
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-01PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..501
PCT-US96-01471-1

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x PCT-US96-01471-1/rev ..
Align seg 1/1 to reverse of: PCT-US96-01471-1 from: 1 to: 504
1 GluAlaGly*****Ser 6
|||||
361 GAGCGAGGAGCAGCTCT 344
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-581-528A-5

seq_documentation_block:
; Sequence 5, Application US/08581528A
; Patent No. 5986058
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; APPLICANT: Huvnh, Thanh
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-7
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,528A
; FILING DATE: 03-Sept-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,670
; FILING DATE: 09-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lisa A. Haile, Ph.D.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/081001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
```

```
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: GDF-7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..516
US-08-581-528A-5

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-581-528A-5/rev ..
Align seg 1/1 to reverse of: US-08-581-528A-5 from: 1 to: 519
1 GluAlaGly*****Ser 6
|||||
404 GAGCGTGGCGCAGCGTCG 387
seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-07799-5

seq_documentation_block:
; Sequence 5, Application PC/TUS9407799
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-7
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07799
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN, LISA A., PH.D.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: FD-2348
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: GDF-7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..516
```

PCT-US94-07799-5

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x PCT-US94-07799-5/rev ..
Align seg 1/1 to reverse of: PCT-US94-07799-5 from: 1 to: 519

1 GluAlaGly*****Ser 6

|||||

404 GAGCGTGGCGAGCGTCG 387

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-292-345B-1

seq_documentation_block:

; Sequence 1, Application US/08292345B

; Patent No. 6001968

; GENERAL INFORMATION:

; APPLICANT: THE ROCKEFELLER UNIVERSITY

; TITLE OF INVENTION: MODULATORS OF BODY WEIGHT, CORRESPONDING NUCLEIC

; TITLE OF INVENTION: ACIDS AND PROTEINS, AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/292,345B

; FILING DATE: August 17, 1994

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-087

; TELEPHONE: 201 487-5800

; TELEFAX: 201 343-1684

; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 701 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Mus musculus

US-08-292-345B-1

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x US-08-292-345B-1/rev ..

Align seg 1/1 to reverse of: US-08-292-345B-1 from: 1 to: 701

1 GluAlaGly*****Ser 6

|||||

407 GAGCAGGGAGCAGCTCT 390

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-003-081-1

seq_documentation_block:

; Sequence 1, Application US/09003081

; Patent No. 5968779

; GENERAL INFORMATION:

; APPLICANT: Campfield, Arthur Dr.

; APPLICANT: Devos, Rene Dr.

; APPLICANT: Guisez, Yves Dr.

; TITLE OF INVENTION: Recombinant Obese (OB) Proteins

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann-La Roche, Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/003,081

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA: 08/435,777

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Picut, Catherine A

; REGISTRATION NUMBER: 37419

; REFERENCE/DOCKET NUMBER: 9165

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-4387

; TELEFAX: (201) 235-2363

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 702 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-09-003-081-1

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x US-09-003-081-1/rev ..

Align seg 1/1 to reverse of: US-09-003-081-1 from: 1 to: 702

1 GluAlaGly*****Ser 6

|||||

396 GAGCAGGGAGCAGCTCT 379

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-648-262-1

seq_documentation_block:

; Sequence 1, Application US/08648262

; Patent No. 6025324

```

; GENERAL INFORMATION:
; APPLICANT: Bailon, Pascal Mr.
; APPLICANT: Campfield, Arthur Dr.
; APPLICANT: Devos, Rene Dr.
; APPLICANT: Guisez, Yves Dr.
; TITLE OF INVENTION: Pegylated obese (OB) Proteins
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche, Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/648,262
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Epstein, William H.
; REGISTRATION NUMBER: 20008
; REFERENCE/DOCKET NUMBER: 9281
; TELEPHONE: (201) 235-3723
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-648-262-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-648-262-1/rev ..
Align seg 1/1 to reverse of: US-08-648-262-1 from: 1 to: 702

1 GluAlaGly*****Ser 6
|||||
396 GAGCGAGGAGCAGCTCT 379

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-648-263-1

seq_documentation_block:
; Sequence 1, Application US/08648263
; Patent No. 6025325
; GENERAL INFORMATION:
; APPLICANT: Campfield, Arthur
; APPLICANT: Devos, Rene
; APPLICANT: Guisez, Yves
; TITLE OF INVENTION: RECOMBINANT OBESE (OB) PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: USA

```

```

; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/648,263
; FILING DATE: 15-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,629
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,777
; FILING DATE: 05-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J.
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4105/175-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)235-4387
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-648-263-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-648-263-1/rev ..
Align seg 1/1 to reverse of: US-08-648-263-1 from: 1 to: 702

1 GluAlaGly*****Ser 6
|||||
396 GAGCGAGGAGCAGCTCT 379

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-856-253-3

seq_documentation_block:
; Sequence 3, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```



```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA: US 60/017,678
; APPLICATION NUMBER: 16-MAY-1996
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 849 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-856-253-3

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-856-253-3 ..
Align seg 1/1 to: US-08-856-253-3 from: 1 to: 849

1 GluAlaGly*****Ser 6
|||||
316 GAAGCGGGAACAAGTAGT 333

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-936-165A-83

seq_documentation_block:
; Sequence 83, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
```

```
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 924 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-83

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-936-165A-83 ..
Align seg 1/1 to: US-08-936-165A-83 from: 1 to: 924

1 GluAlaGly*****Ser 6
|||||
826 GAAGCTGGTGCTACATCA 843

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-362-670B-1

seq_documentation_block:
; Sequence 1, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 926 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
;   IMMEDIATE SOURCE:
;   CLONE: vl-1
;   FEATURE:
;     NAME/KEY: mat_peptide
;     LOCATION: 571..882
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..882
; US-08-362-670B-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-362-670B-1/rev ..

  Align seg 1/1 to reverse of: US-08-362-670B-1 from: 1 to: 926

      1 GluAlaGly*****Ser 6
      |||||||
      770 GAGCGCGCGCGCGTCT 753

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-333-576C-1

seq_documentation_block:
; Sequence 1, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
;   APPLICANT: Celeste, Anthony J.
;   APPLICANT: Wozney, John
;   APPLICANT: Rosen, Vicki A.
;   APPLICANT: Wolfman, Neil
;   APPLICANT: Thomsen, Gerald H.
;   APPLICANT: Melton, Douglas A.
;   TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
;   NUMBER OF SEQUENCES: 37
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: GENETICS INSTITUTE, INC.
;     STREET: 87 Cambridgepark Drive
;     CITY: Cambridge
;     STATE: Massachusetts
;     COUNTRY: USA
;     ZIP: 02140
;   COMPUTER READABLE FORM:
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/333,576C
;       FILING DATE: No. 6027919ember 2, 1994
;       CLASSIFICATION: 435
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Lazar, Steven R.
;       REGISTRATION NUMBER: 32,618
;       REFERENCE/DOCKET NUMBER: 5202-B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 617 498-8260
;       TELEFAX: 617 876-5851
;     INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 926 base pairs
;         TYPE: nucleic acid
```

```
; LENGTH: 926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: vl-1
; FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: 571..882
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..882
; US-08-333-576C-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-333-576C-1/rev ..

  Align seg 1/1 to reverse of: US-08-333-576C-1 from: 1 to: 926

      1 GluAlaGly*****Ser 6
      |||||||
      770 GAGCGCGCGCGCGTCT 753

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-1

seq_documentation_block:
; Sequence 1, Application US/08808324
; Patent No. 6284872
; GENERAL INFORMATION:
;   APPLICANT: Celeste, Anthony J.
;   APPLICANT: Wozney, John
;   APPLICANT: Rosen, Vicki A.
;   APPLICANT: Wolfman, Neil
;   APPLICANT: Thomsen, Gerald H.
;   APPLICANT: Melton, Douglas A.
;   TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: GENETICS INSTITUTE, INC.
;     STREET: 87 Cambridgepark Drive
;     CITY: Cambridge
;     STATE: Massachusetts
;     COUNTRY: USA
;     ZIP: 02140
;   COMPUTER READABLE FORM:
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/808,324
;       FILING DATE: Herewith
;       CLASSIFICATION: 514
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Lazar, Steven R.
;       REGISTRATION NUMBER: 32,618
;       REFERENCE/DOCKET NUMBER: 5202-D
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 617 498-8260
;       TELEFAX: 617 876-5851
;     INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 926 base pairs
;         TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: v1-1
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 571..882
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..882
US-08-808-324-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-808-324-1/rev ..
  Align seg 1/1 to reverse of: US-08-808-324-1 from: 1 to: 926

  1 GluAlaGly*****Ser 6
  |||||
  770 GAGGCGGCGCGCGTCT 753

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-14030A-1

seq_documentation_block:
; Sequence 1, Application PCT/US94/14030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: v1-1
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 571..882
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..882
PCT-US94-14030A-1

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x PCT-US94-14030A-1/rev ..
  Align seg 1/1 to reverse of: PCT-US94-14030A-1 from: 1 to: 926

  1 GluAlaGly*****Ser 6
  |||||
  770 GAGGCGGCGCGCGTCT 753

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-581-148C-10

seq_documentation_block:
; Sequence 10, Application US/08581148C
; Patent No. 6060644
; GENERAL INFORMATION:
; APPLICANT: Schnable, Patrick S.
; APPLICANT: Robertson, Donald S.
; APPLICANT: Hansen, Joel D.
; APPLICANT: Nikolau, Basil J.
; APPLICANT: Xu, Xiaojie
; APPLICANT: Xia, Yiji
; TITLE OF INVENTION: ISOLATION AND USE OF CUTICULAR LIPID
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,148C
; FILING DATE: 29-DEC-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Larcher, Carol
; REGISTRATION NUMBER: 35243
; REFERENCE/DOCKET NUMBER: 71380
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-581-148C-10

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-581-148C-10/rev ..
Align seg 1/1 to reverse of: US-08-581-148C-10 from: 1 to: 950

1 GluAlaGly*****Ser 6
|||||
870 GAGCGCGGAGACCTCG 853

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5202428-1
seq_documentation_block:
; Patent No. 5202428
; APPLICANT: SCHUBERT, DAVID
; TITLE OF INVENTION: DNA ENCODING NEUROTROPHIC GROWTH FACTOR
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/590,359
; FILING DATE: 27-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 541,276
; FILING DATE: 20-JUN-1990
; SEQ ID NO:1
; LENGTH: 1159
5202428-1

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x 5202428-1/rev ..
Align seg 1/1 to reverse of: 5202428-1 from: 1 to: 1159

1 GluAlaGly*****Ser 6
|||||
46 GAGCGTGGAGTCCAGC 29

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-746-397-1
seq_documentation_block:
; Sequence 1, Application US/08746397
; Patent No. 6130061
; GENERAL INFORMATION:
; APPLICANT: NI, ET AL.
; TITLE OF INVENTION: Human Stem Cell Antigen 2
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,397
; FILING DATE: 11/8/96
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/007,287
; FILING DATE: 11/9/95
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-494
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
US-08-746-397-1

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-746-397-1/rev ..
Align seg 1/1 to reverse of: US-08-746-397-1 from: 1 to: 1163

1 GluAlaGly*****Ser 6
|||||
80 GAGCAGGTACCAGCAGC 63

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-362-670B-29
seq_documentation_block:
; Sequence 29, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: murine MV1
; NAME/KEY: CDS
; LOCATION: 2..721
; US-08-362-670B-29

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-362-670B-29/rev ..
  Align seg 1/1 to reverse of: US-08-362-670B-29 from: 1 to: 1203

      1 GluAlaGly*****Ser 6
      |||||
609 GAGGCTGGCGCAGCGTCG 592

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-333-576C-29

seq_documentation_block:
; Sequence 29, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,576C
; FILING DATE: No. 6027919ember 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: murine MV1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..721
; US-08-333-576C-29

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM4 x US-08-333-576C-29/rev ..
  Align seg 1/1 to reverse of: US-08-333-576C-29 from: 1 to: 1203

      1 GluAlaGly*****Ser 6
      |||||
609 GAGGCTGGCGCAGCGTCG 592

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-29

seq_documentation_block:
; Sequence 29, Application US/08808324
; Patent No. 6284872
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,324
; FILING DATE: Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: murine MV1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..721
```

US-08-808-324-29

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x US-08-808-324-29/rev ..

Align seg 1/1 to reverse of: US-08-808-324-29 from: 1 to: 1203

1 GluAlaGly*****Ser 6

|||||
609 GAGGCTGGCGCAGCGTCG 592

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-14030A-29

seq_documentation_block:
; Sequence 29, Application PCT/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:

ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/14030A
FILING DATE: Herewith

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/164,103
FILING DATE: 07-DEC-1993
APPLICATION NUMBER: US 08/217,780
FILING DATE: 25-MAR-1994
APPLICATION NUMBER: US 08/333,576
FILING DATE: 02-NOV-1994

ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202D-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5951

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1203 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: murine MV1
FEATURE:

NAME/KEY: CDS
LOCATION: 2..721
PCT-US94-14030A-29

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x PCT-US94-14030A-29/rev ..

Align seg 1/1 to reverse of: PCT-US94-14030A-29 from: 1 to: 1203

1 GluAlaGly*****Ser 6

|||||
609 GAGGCTGGCGCAGCGTCG 592

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-362-670B-27

seq_documentation_block:
; Sequence 27, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:

ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,670B
FILING DATE: December 22, 1994

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5951

INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1233 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: DNA encoding BMP2 propeptide/BMP-12 mature
FEATURE:

NAME/KEY: CDS
LOCATION: 1..1233
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 847..1233
US-08-362-670B-27

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```
alignment_block:
BASK-853-CLAIM4 x US-08-362-670B-27/rev ..
Align seg 1/1 to reverse of: US-03-362-670B-27 from: 1 to: 1233

1 GluAlaGly*****Ser 6
|||||
1121 GAGCGCGCGCGGTCT 1104

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-333-576C-27

seq_documentation_block:
; Sequence 27, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,576C
; FILING DATE: No. 6027919ember 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: DNA encoding BMP2 propeptide/BMP-12 mature
; CLONE: peptide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1233
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 847..1233
; US-08-333-576C-27

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-333-576C-27/rev ..
```

```
Align seg 1/1 to reverse of: US-08-333-576C-27 from: 1 to: 1233

1 GluAlaGly*****Ser 6
|||||
1121 GAGCGCGCGCGGTCT 1104

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-808-324-27

seq_documentation_block:
; Sequence 27, Application US/08808324
; Patent No. 6284872
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,324
; FILING DATE: Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: DNA encoding BMP2 propeptide/BMP-12 mature
; CLONE: peptide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1233
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 847..1233
; US-08-808-324-27

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-808-324-27/rev ..
Align seg 1/1 to reverse of: US-08-808-324-27 from: 1 to: 1233

1 GluAlaGly*****Ser 6
```

```
|||||
1121 GAGCGCGCGCGGTCT 1104

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_94-14030A-27

seq_documentation_block:
; Sequence 27, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: DNA encoding BMP2 propeptide/BMP-12 mature
; CLONE: peptide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1233
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 847..1233
PCT-US94-14030A-27

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x PCT-US94-14030A-27/rev ..
Align seg 1/1 to reverse of: PCT-US94-14030A-27 from: 1 to: 1233
1 GluAlaGly*****Ser 6
|||||

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-945-296-1

seq_documentation_block:
; Sequence 1, Application US/08945296
; Patent No. 5972609
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; APPLICANT: Dennis, Carina
; TITLE OF INVENTION: Utrrophin gene promoter
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5972609th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,296
; FILING DATE: 22-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00979
; FILING DATE: 24-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9508236.8
; FILING DATE: 24-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9526028.7
; FILING DATE: 20-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1246 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-945-296-1

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-08-945-296-1 ..
Align seg 1/1 to: US-08-945-296-1 from: 1 to: 1246
1 GluAlaGly*****Ser 6
|||||
925 GAAGCGGCGCAGCAGC 942

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-405-112-1

seq_documentation_block:
; Sequence 1, Application US/09405112
; Patent No. 6087111
; GENERAL INFORMATION:
```



```
;
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; APPLICANT: Dennis, Carina
; TITLE OF INVENTION: Utrrophin gene promoter
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 NO. 6087111th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/405,112
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,296
; FILING DATE: 22-OCT-1997
; APPLICATION NUMBER: PCT/GB96/00979
; FILING DATE: 24-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9508236.8
; FILING DATE: 24-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9526028.7
; FILING DATE: 20-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1246 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-405-112-1

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x US-09-405-112-1 ..
Align seg 1/1 to: US-09-405-112-1 from: 1 to: 1246

1 GluAlaGly*****Ser 6
|||||
925 GAAGCGGCAGCAGCAGC 942

seq_name: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:PCT-US91-02766-19

seq_documentation_block:
; Sequence 19, Application PC/TUS9102766
; GENERAL INFORMATION:
; APPLICANT: NASCIMENTO, CARLOS G
; APPLICANT: CALDERON-CACIA, MARIA D.
; TITLE OF INVENTION: GLYCOSYLATED PDGF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Irell & Manella
; STREET: 545 Middlefield Road, Suite 200
```

```
;
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02766
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/515,474
; FILING DATE: 26-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 2300-0105.40
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1247 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 396..983
; PCT-US91-02766-19

alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x PCT-US91-02766-19/rev ..
Align seg 1/1 to reverse of: PCT-US91-02766-19 from: 1 to: 1247

1 GluAlaGly*****Ser 6
|||||
222 GAGCGAGCGCAGCGCCAGC 205
```

THIS PAGE BLANK (USPTO)

OM of: BASK-853-CLAIM4 to: EST:* out_format : pfs

Date: Sep 24, 2002 12:52 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-O=/cgn2.1/USPTO_spool/BASKAR853079/runat_17092002_103407_1808/app_query.fasta_1.98
-DB=EST -QFWT=fastap -SUFFIX=mod.ist -GAPOP=12.000 -CAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-CAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blcsun62 -TRANS=human40.cdi
-LIST=50 -DLOCAL=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=BASKAR853079 -SCGNL_1.3956
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: BASK-853-CLAIM4

Query length: 6

Database: EST:*

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 2564.800000

score_list:

Sequence	Strd	Orig	Zscore	EScore	Len	Documentation
gb_est2:B6034279	+	19.00	103.40	1.6e+04	64	! BG034279 602302354F1 NIH_MGC_87
gb_est1:AA242543	-	19.00	102.91	1.7e+04	68	! AA242543 ESTKB010 4XF1/106/KO15
gb_est1:AL633696	-	19.00	102.32	1.9e+04	73	! AL633696 AL633696 XGP-gastrula.8
gb_est2:BG574341	+	19.00	101.99	2.0e+04	76	! BG574341 602596265F1 NIH_MGC_87
gb_est2:TA9595	+	19.00	101.67	2.0e+04	79	! T89595 yd98ell.r1 Soares fetal.1
gb_est2:BG31927	+	19.00	101.57	2.1e+04	80	! BG31927 602807326F1 NCI_CGAP_BH
gb_est2:BG392056	+	19.00	101.17	2.2e+04	84	! BG392056 602410021F1 NIH_MGC_92
gb_est2:BG1086268	+	19.00	101.17	2.2e+04	84	! BG1086268 602849676F1 NIH_MGC_10
gb_est2:T61780	+	19.00	100.79	2.3e+04	88	! T61780 yb93h01.r1 Stratagene.11v
gb_est2:BF796213	+	19.00	100.34	2.4e+04	93	! BF796213 602260191F1 NIH_MGC_85
gb_gss:BHG625950	-	19.00	99.82	2.6e+04	99	! BHG625950 1007110005.2EL.x1 1007
gb_est2:BG171437	+	19.00	99.74	2.6e+04	100	! BG171437 602321943F1 NIH_MGC_89
gb_est2:BG102440	+	19.00	99.50	2.7e+04	103	! BG102440 RH122_24_H12.B1_A003.R
gb_est2:BG244884	+	19.00	99.42	2.7e+04	104	! BG244884 TCBAP1E2768 Pediatric
gb_est2:BF823134	+	19.00	99.11	2.8e+04	108	! BF823134 MRL-RT0079-151200-003-
gb_est1:AT307952	+	19.00	99.03	2.9e+04	109	! AT307952 ta86g01.x1 NCI_CGAP_Bd
gb_est1:AT547791	+	19.00	99.03	2.9e+04	109	! AT547791 UI-R-C3-sj-h-04-0-UI-S
gb_est2:BF175147	+	19.00	98.96	2.9e+04	110	! BF175147 MYE498 Myeloma (MYE)
gb_gss:AZ882861	-	19.00	98.81	2.9e+04	112	! AZ882861 RPCI-23-186P2.TV RPCI-1
gb_gss:BH406737	-	19.00	98.81	2.9e+04	112	! BH406737 QV4-HT0538-170300-142-
gb_gss:TA157B06P	+	19.00	98.81	2.9e+04	112	! TA157B06P T. brucei sheared gene
gb_est1:AF090252	+	19.00	98.67	3.0e+04	114	! AF090252 AF090252 Rat limonene
gb_est1:AV624809	+	19.00	98.59	3.0e+04	115	! AV624809 AV624809 Chlamydomonas
gb_est1:BE170537	+	19.00	98.45	3.1e+04	117	! BE170537 QV4-HT0538-170300-142-
gb_gss:BH096312	-	19.00	98.45	3.1e+04	117	! BH096312 RPCI-24-391J2.TV RPCI-1
gb_gss:AZ442930	-	19.00	98.38	3.1e+04	118	! AZ442930 LM0237D24F Mouse 10kb
gb_est2:BM434428	+	19.00	98.31	3.1e+04	119	! BM434428 1RT10E12 Bos taurus Re
gb_gss:AZ256123	-	19.00	98.31	3.1e+04	119	! AZ256123 RPCI-23-100B21.TV RPCI
gb_est1:AT014429	+	19.00	98.18	3.2e+04	121	! AT014429 993f07.s1 NCI_CGAP_R
gb_est2:BG032062	+	19.00	98.11	3.2e+04	122	! BG032062 602300576F1 NIH_MGC_87
gb_est1:AW880397	+	19.00	97.91	3.3e+04	125	! AW880397 QV0-CT0030-100400-188-
gb_est2:BG383537	+	19.00	97.91	3.3e+04	125	! BG383537 301674 MARC 1P1G Sus s
gb_est2:BI779080	+	19.00	97.91	3.3e+04	125	! BI779080 EBR001_S0002_N24_R IGH
gb_est1:AW97381	+	19.00	97.72	3.4e+04	128	! AW97381 RCO-BN0050-160200-011-
gb_est2:BF983162	+	19.00	97.72	3.4e+04	128	! BF983162 602309067F1 NIH_MGC_88
gb_est2:BG253687	+	19.00	97.65	3.4e+04	129	! BG253687 602366742F1 NIH_MGC_91
gb_est2:BM445566	-	19.00	97.65	3.4e+04	129	! BM445566 11L3A7.ab1 Bos taurus
gb_est2:BG166289	+	19.00	97.53	3.5e+04	131	! BG166289 602345480F1 NIH_MGC_89
gb_est2:BG258468	+	19.00	97.53	3.5e+04	131	! BG258468 602379947F1 NIH_MGC_92

gb_est2:B6034748	+	19.00	97.53	3.5e+04	131	! BG034748 602419577F1 NIH_MGC
gb_est1:AA242543	+	19.00	97.40	3.5e+04	133	! AA242543 mx12d05.r1 Soares m
gb_gss:AZ909098	+	19.00	97.40	3.5e+04	133	! AZ909098 RPCI-24-191G10.TV R
gb_est1:AW840673	-	19.00	97.34	3.6e+04	134	! AW840673 RC4-CN0007-090300-0
gb_est2:BF903294	+	19.00	97.34	3.6e+04	134	! BF903294 PM4-WT0201-181200-0
gb_est2:BI856341	+	19.00	97.34	3.6e+04	134	! BI856341 603387591F1 NIH_MGC
gb_gss:BG388955	+	19.00	97.28	3.6e+04	135	! BG388955 602414709F1 NIH_MGC
gb_est2:BI090037	+	19.00	97.28	3.6e+04	135	! BI090037 602857337F1 NIH_MGC
gb_est1:AW840788	-	19.00	97.22	3.6e+04	136	! AW840788 RC4-CN0007-090300-0
gb_est1:AW922200	+	19.00	97.22	3.6e+04	136	! AW922200 LG1_346_H08.B1_A002

seq_name: gb_est2:BG034279

seq_documentation_block:

LOCUS BG034279 64 bp mRNA linear EST 24-JAN-2001
DEFINITION 602302354F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4403869 5',
mRNA sequence.

ACCESSION BG034279

VERSION BG034279.1 GI:12427431

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 64)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLMI0114 row: 1 column: 14

High quality sequence stop: 64.

Location/Qualifiers

1..64

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4403869"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/notes="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-gr primed.

Average insert size 1.383 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: This is a NIH_MGC Library."

BASE COUNT 9 a 23 c 26 g 6 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BG034279

Align seg 1/1 to: BG034279 from: 1 to: 64

1 GluAlaGly*****Ser 6

|||||||

38 GAAGCCGCCGCGTTCG 55

seq_name: gb_est1:AA566112

seq_documentation_block:

```

LOCUS      AA566112                68 bp    mRNA    linear    EST 01-JUN-1998
DEFINITION ESTKB010 4AF1/106/KO15 library (Lap-Chee Tsui) Homo sapiens CDNA
clone 5.00E+35, mRNA sequence.
ACCESSION  AA566112
VERSION     AA566112.1  GI:3171685
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 68)
AUTHORS   Torigoe,K., Harada,T., Kusaba,H., Uchiyumi,T., Kohno,K., Green,E.D.,
            Scherer,S.W., Tsui,L.C., Schlessinger,D., Kuwano,M. and Wada,M.
TITLE     Localization of 67 exons on a YAC contig spanning 1.5 Mb around the
            multidrug resistance gene region of human chromosome 7q21.1
JOURNAL   Genomics 49 (1), 14-22 (1998)
MEDLINE   98234539
COMMENT   Contact: Wada M
            Department of Biochemistry
            Kyushu University School of Medicine
            Maidashi Fukuoka 812-82, Japan
            Tel: 8 1-92-642-6100
            Fax: 81-92-642-6203
            Email: wada@mailserver.med.kyushu-u.ac.jp
            This EST was isolated by exon amplification from HSC7E515
            Insert Length: 68 Std Error: 0.00
Seq primer: SP6
            Location/Qualifiers
            1..68
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="5.00E+35"
            /clone_lib="4AF1/106/KO15 library (Lap-Chee Tsui)"
            /cell_line="4AF1/106/KO15"
            /notes="Chromosome 7 specific library containing human DNA
            derived from the somatic cell hybrid cell line
            4AF1/106/KO15"
BASE COUNT  18 a 14 c 24 g 12 t
ORIGIN
alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167       Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
    BASK-853-CLAIM4 x AA566112 ..
Align seg 1/1 to: AA566112 from: 1 to: 68
    1 GluAlaGly*****Ser 6
      |||||||
      46 GAAGCAGGCACATCTTCA 63
seq_name: gb_est1:AL633696
seq_documentation_block:
LOCUS      AL633696                73 bp    mRNA    linear    EST 02-NOV-2001
DEFINITION AL633696 XGC-gastrula Silurana tropicalis cDNA clone TGas017021 5',
mRNA sequence.
ACCESSION  AL633696
VERSION     AL633696.1  GI:16603179
KEYWORDS   EST.
SOURCE     western clawed frog.
ORGANISM   Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopodinae; Silurana.
            1 (bases 1 to 73)
REFERENCE  1 (bases 1 to 73)
AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE     Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL   Unpublished (2001)

```

```

COMMENT   Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas017021.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            Location/Qualifiers
            1..73
            /organism="Silurana tropicalis"
            /db_xref="taxon:8364"
            /clone="TGas017021"
            /clone_lib="XGC-gastrula"
            /dev_stage="gastrula (stages 10.5-13 mixed)"
            /lab_host="Escherichia coli XL1-blue"
            /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
            was oligo dt primed from 5ug of poly A+ RNA from stages
            10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
            into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end."
BASE COUNT  25 a 17 c 21 g 10 t
ORIGIN
alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167       Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
    BASK-853-CLAIM4 x AL633696/rev ..
Align seg 1/1 to reverse of: AL633696 from: 1 to: 73
    1 GluAlaGly*****Ser 6
      |||||||
      43 GAAGCAGGGCGCTCGTCA 26
seq_name: gb_est2:BG574341
seq_documentation_block:
LOCUS      BG574341                76 bp    mRNA    linear    EST 10-APR-2001
DEFINITION 602596265F1 NIH_MGC_87 Homo sapiens CDNA clone IMAGE:4705048 5',
mRNA sequence.
ACCESSION  BG574341
VERSION     BG574341.1  GI:13581994
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 76)
REFERENCE  1 (bases 1 to 76)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue procurement: DCTD/DTP
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            plate: LLAM10571 row: a column: 17
            High quality sequence stop: 76.
            Location/Qualifiers
            1..76
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4705048"

```

/clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 13 a 26 c 30 g 7 t
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 BASK-853-CLAIM4 x BG574341 ..

Align seg 1/1 to: BG574341 from: 1 to: 76

1 GluAlaGly*****Ser 6
 |||||
 50 GAAGCGCGCGCCAGTGC 67

seq_name: gb_est2:T89595

seq_documentation_block:

LOCUS T89595 79 bp mRNA linear EST 20-MAR-1995
 DEFINITION Y98a11.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
 IMAGE:116300 5' similar to SP:ZK637.5 CE00436 ARSA ;, mRNA
 sequence.

ACCESSION T89595
 VERSION T89595.1 GI:718108
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
 M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
 R., Williamson,A., Wohlmann,P. and Wilson,R.
 The WashU-Merck EST Project

TITLE

JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 623

High quality sequence starts: 1 High quality sequence stops: 1

Source: IMAGE Consortium LLNL This clone is available royalty-free
 through LLNL; contact the IMAGE Consortium (info@image.llnl.gov)

for further information! Trace considered overall poor quality

Possible reversed clone; similarity on wrong strand

Insert Length: 623 Std Error: 0.00

Seq primer: M13Rpl

High quality sequence stop: 1.

FEATURES

source

1. .79 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="GDB:471917"

/db_xref="taxon:9606"

/clone="IMAGE:116300"

/clone_lib="Soares fetal liver spleen INFLS"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: liver and Spleen; Vector: pT73D (Pharmacia)

with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5' AACTGGAAGATTAATAAGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 15 a 24 c 20 g 16 t 4 others
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 BASK-853-CLAIM4 x T89595 ..

Align seg 1/1 to: T89595 from: 1 to: 79

1 GluAlaGly*****Ser 6
 |||||
 61 GAGGCTGGCTCACTCC 78

seq_name: gb_est2:BG912927

seq_documentation_block:

LOCUS BG912927 80 bp mRNA linear EST 05-JUN-2001
 DEFINITION G02807376F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939351
 5', mRNA sequence.

ACCESSION BG912927
 VERSION BG912927.1 GI:14293403
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps@email.nih.gov

Tissue Procurement: David N. Louis, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10876 row: d column: 08

High quality sequence stop: 80.

FEATURES

source

1. .80 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4939351"

/clone_lib="NCI_CGAP_Brn67"

/tissue_type="anaplastic oligodendroglioma with lp/19q

loss"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.3 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 14 a 28 c 31 g 7 t
 ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BG912927 ..

Align seg 1/1 to: BG912927 from: 1 to: 80

1 GluAlaGly*****Ser 6

|||||
54 GAAGCGCGCGCCAGTTCG 71

seq_name: gb_est2:BG392056

seq_documentation_block:

LOCUS BG392056 84 bp mRNA linear EST 12-MAR-2001
DEFINITION 602410021F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4538597 5',
mRNA sequence.

ACCESSION BG392056

VERSION BG392056.1 GI:13285504

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 84)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ARCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10465 row: j column: 06

High quality sequence stop: 84.

FEATURES

source

1..84

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4538597"

/clone_lib="NIH_MGC_92"

/tissue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 Kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

14 a 30 c 33 g 7 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BG392056 ..

Align seg 1/1 to: BG392056 from: 1 to: 84

1 GluAlaGly*****Ser 6

|||||
58 GAAGCGCGCGCCAGTTCG 75

seq_name: gb_est2:BI086268

seq_documentation_block:

LOCUS BI086268 84 bp mRNA linear EST 20-JUN-2001
DEFINITION 602849676F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4991275 5',
mRNA sequence.

ACCESSION BI086268

VERSION BI086268.1 GI:14504598

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 84)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ARCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11008 row: g column: 20

High quality sequence stop: 84.

FEATURES

source

1..84

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4991275"

/clone_lib="NIH_MGC_10"

/cell_line="MGC36"

/lab_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.5 Kb. Library prepared by Life

Technologies."

16 a 26 c 33 g 9 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BI086268 ..

Align seg 1/1 to: BI086268 from: 1 to: 84

1 GluAlaGly*****Ser 6

|||||
58 GAAGCGCGCGCCAGTTCG 75

seq_name: gb_est2:BI086268

seq_documentation_block:

LOCUS BI086268 84 bp mRNA linear EST 14-FEB-1995
DEFINITION YB93h01.r1 Stratagene liver (#937224) Homo sapiens cDNA clone
IMAGE:78769 5', mRNA sequence.

ACCESSION T61780

VERSION T61780

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 88)

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,

Chissoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins

,M., Hultman,M., Kucaba,T., Lacy,W., Le.N., Mardis,E., Moore

B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)

MEDLINE

97044478

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1071

High quality sequence stops: 54 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1071 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 54.

FEATURES

source

1. .88
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDE:498514"
/db_xref="taxon:9606"
/clone="IMAGE:78769"
/clone_lib="Stratagene liver (#937224)"
/sex="male"
/dev_stage="49 years old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: liver; Vector: pBluescript SK; Site.1: EcoRI
; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Hepatectomy from normal male caucasian. Average insert
size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTGTGTTTTTTTTTTT 3'"

BASE COUNT 22 a 24 c 28 g 13 t 1 others

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x T61780 ..

Align seg 1/1 to: T61780 from: 1 to: 88

1 GluAlaGly*****Ser 6

|||||
51 GAAGCAGGAGCATCTCT 68

seq_name: gb_est2:BF796213

seq_documentation_block:

LOCUS BF796213 93 bp mRNA linear EST 12-JAN-2001
DEFINITION 602260191F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4343422 5',
mRNA sequence.

ACCESSION BF796213

VERSION BF796213.1 GI:12101267

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 93)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapb-r@mail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9960 row: e column: 23

High quality sequence stop: 93.

FEATURES

source

1. .93
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4343422"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 15 a 36 c 34 g 8 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BF796213 ..

Align seg 1/1 to: BF796213 from: 1 to: 93

1 GluAlaGly*****Ser 6

|||||
67 GAAGCGGGCGCGAGTTCG 84

seq_name: gb_gss:BH625950

seq_documentation_block:

LOCUS BH625950 99 bp DNA linear GSS 30-JAN-2002
DEFINITION 1007110D05.2EL_x1 1007 - RescueMu Grid H Zea mays genomic, DNA

ACCESSION BH625950

VERSION BH625950.1 GI:18439201

KEYWORDS GSS.

SOURCE Zea mays.

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 99)

Walbot, V.

Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1007110 column: 11

Class: transposon-tagged.

FEATURES

source

1. .99
Location/Qualifiers

/organism="Zea mays"

/cultivar="mixed background W23/Al88/B73"

/db_xref="taxon:4577"

```

/clone_lib="1007 - RescueMu Grid H"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site.1: BamHI; Site.2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT      20 a  40 c  18 g  21 t
ORIGIN

```

```

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:
BASK-853-CLAIM4 x BH625950/rev ..

```

```
Align seg 1/1 to reverse of: BH625950 from: 1 to: 99

```

```

1 GluAlaGly*****Ser 6
|||||
19 GAGGCTGGAACGGCAGC 2

```

```
seq_name: gb_est2:BG171437

```

```

seq_documentation_block:
LOCUS      BG171437                100 bp  mRNA  linear  EST 06-FEB-2001
DEFINITION 602321943F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4425202 5',
            mRNA sequence.
ACCESSION  BG171437
VERSION    BG171437.1 GI:12678140
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 100)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-re@mail.nih.gov

```

```

Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10170 row: e column: 11
High quality sequence start: 3
High quality sequence stop: 100.

```

```

FEATURES             Location/Qualifiers
     source           1..100
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:4425202"
     /clone_lib="NIH_MGC_89"
     /tissue_type="hypernephroma, cell line"
     /lab_host="DH10B (phage-resistant)"
     /note="Organ: kidney; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed."

```

```

Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT      17 a  36 c  35 g  12 t
ORIGIN

```

```

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:
BASK-853-CLAIM4 x BG171437 ..

```

```
Align seg 1/1 to: BG171437 from: 1 to: 100

```

```

1 GluAlaGly*****Ser 6
|||||
74 GAAGCGGGCGCAGTTCG 91

```

```
seq_name: gb_est2:BG102440

```

```
seq_documentation_block:
LOCUS      BG102440                103 bp  mRNA  linear  EST 30-JAN-2001
DEFINITION RHIZ2_24_HI2.bl_A003 Rhizome2 (RHIZ2) Sorghum propinquum cDNA, mRNA
            sequence.
ACCESSION  BG102440
VERSION    BG102440.1 GI:12617273
KEYWORDS   EST.
SOURCE     Sorghum propinquum.
ORGANISM   Sorghum propinquum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE  1 (bases 1 to 103)
            Cordonnier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt
            ,L.H.
            An EST database from Sorghum: Sorghum propinquum rhizomes
            Unpublished (2000)
            Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Seq primer: JEN REV
            High quality sequence stop: 102
            POLYA-No.

```

```

FEATURES             Location/Qualifiers
     source           1..103
     /organism="Sorghum propinquum"
     /db_xref="taxon:132711"
     /clone_lib="Rhizome2 (RHIZ2)"
     /note="Organ: Rhizomes; Vector: pBluescript II from lambda
zap II; Site.1: XhoI; Site.2: EcoRI; The library was made
from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."

```

```

BASE COUNT      14 a  48 c  23 g  18 t
ORIGIN

```

```

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:
BASK-853-CLAIM4 x BG102440 ..

```

```
Align seg 1/1 to: BG102440 from: 1 to: 103

```


1 GluAlaGly*****Ser 6
 |||||
 5 GAGGCTGGACGAGCTCT 22

seq_name: gb_est2:BE244884

seq_documentation_block:
 LOCUS BE244884 104 bp mRNA linear EST 03-OCT-2001
 DEFINITION TCBAP1E2768 Pediatric pre-B cell acute lymphoblastic leukemia
 Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP2768, mRNA
 sequence.
 ACCESSION BE244884
 VERSION BE244884.1 GI:9096714
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 104)
 AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R., Muzny,D.,
 Bouck,J.J., Gibbs,R.A. and Margolin,J.F.
 TITLE Pediatric Leukemia cDNA Sequencing Project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@ccc.org
 Citation: Carninci,P. and Hayashizaki,Y. High efficiency
 full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Seq primer: M13 primer.

FEATURES

Location/Qualifiers
 1..104
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCBAP2768"
 /clone_lib="Pediatric pre-B cell acute lymphoblastic
 leukemia Baylor-HGSC project-TCBA"
 /sex="male"
 /tissue_type="Leukopheresis"
 /cell_type="pre-B cell"
 /dev_stage="Pediatric 2 years"
 /lab_host="DH10B"
 /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
 First strand cDNA was primed with an anchored
 XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGCAGAGGAG(T)VN
 3': V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
 was primed with a BamHI-dC primer
 [5'AGAGCTGGATCCGGCCGCAATAATAAT(C) 3'].
 Double-stranded cDNA was then digested with BamHI and XhoI
 and directionally cloned into the BamHI and SalI sites of
 lambda pSB vector. Library was constructed by one round of
 normalization. Library was constructed by Wei Yu at RIKEN
 of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
 Itoh M, Nagakura S, Sasakini, Okazaki Y, Muramatsu M,
 Schneider C, Hayashizaki Y, High efficiency selection of
 full-length cDNA by improved biotinylated cap trapper.,
 DNA Res 4: 1, 51-6, Feb 28, 1997)"
 DNA Res 4: 1, 51-6, Feb 28, 1997)"

BASE COUNT 23 a 18 c 16 t 1 others
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BE244884 ..

Align seg 1/1 to: BE244884 from: 1 to: 104

1 GluAlaGly*****Ser 6
 |||||
 40 GAGGCTGGGCTGCGTCC 57

seq_name: gb_est2:BF823134

seq_documentation_block:
 LOCUS BF823134 108 bp mRNA linear EST 13-JAN-2001
 DEFINITION MR1-RT0079-151200-003-h10 RT0079 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF823134
 VERSION BF823134.1 GI:12163703
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 108)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-RT0079-
 151200-003-h10&t3=2000-12-15&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 78.

FEATURES

Location/Qualifiers
 1..108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RT0079"
 /dev_stage="Adult"
 /note="Organ: kidney_tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 20 a 29 c 29 g 30 t
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BF823134/rev ..

Align seg 1/1 to reverse of: BF823134 from: 1 to: 108

1 GluAlaGly*****Ser 6

72 GAGGACGAGACCGACTCT 55

```
seq_name: gb_est1:AI307952

seq_documentation_block:
LOCUS      AI307952              109 bp      mRNA          linear      EST 18-MAR-1999
DEFINITION ta86g01.x1 NCI_CGAP_Brn20 Homo sapiens cDNA clone IMAGE:2050992 3',
            similar to SW:RS10_HUMAN P46783 40S RIBOSOMAL PROTEIN S10. ;, mRNA
            sequence.
ACCESSION  AI307952
VERSION    AI307952.1  GI:4002587
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 109)
AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGP), Tumor Gene Index
JOURNAL    Unpublished (1998)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Robert Jenkins, M.D., Mark Israel, M.D., Jim
            Jacobson, Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            DNA Sequencing by: Greg Lennon, Ph.D.
            CDNA Library Arrayed by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbr/image/image.html

Trace considered overall poor quality
Insert Length: 163 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 1.

FEATURES
    Source
    1..109
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:2050992"
        /clone_lib="NCI_CGAP_Brn20"
        /tissue_type="oligodendroglioma"
        /dev_stage="adult"
        /lab_host="DH10B"
        /note="Organ: brain; Vector: pAMP1; mRNA made from
            oligodendroglioma tissue, cDNA made by oligo-dT priming.
            Directionally cloned. Size-selected on agarose gel,
            average insert size 500 bp. Primary library,
            non-amplified."
BASE COUNT      30 a      34 c      30 g      15 t
ORIGIN
1
GluAlaGly*****Ser 6
|||||
35 GAGGCTGGGGCTGCCTCA 52

seq_name: gb_est1:AI547791

seq_documentation_block:
LOCUS      AI547791              109 bp      mRNA          linear      EST 22-MAR-1999
DEFINITION UI-R-C3-sj-h-04-0-UI.s1 UI-R-C3 Rattus norvegicus cDNA clone
```

Align seg 1/1 to reverse of: A1547791 from: 1 to: 109

1 GluAlaGly*****Ser 6
|||||

49 GAGGAGGAGCTCTTCC 32

seq_name: gb_est2:BF175147

seq_documentation_block:

LOCUS BF175147 110 bp mRNA linear EST 23-MAR-2001
DEFINITION MYE4498 Myeloma (MYE) cDNA library Homo sapiens cDNA, mRNA
sequence.

ACCESSION BF175147

VERSION BF175147.1 GI:13441361

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 110)

AUTHORS Claudio J.O., Tang H., Khan E.M., Voralia M., Li Z., Cukerman E.,

Francisco-Pabalan O., Liew C.C. and Stewart A.K.

TITLE The transcriptional phenotype of myeloma cells

JOURNAL Unpublished (2000)

COMMENT Contact: A. Keith Stewart, M.D.

Oncology Research

University Health Network

610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada

Tel: (416) 946-4639

Fax: (416) 946-6546

Email: k.stewart@utoronto.ca

PCR Primers

FORWARD: 5'-GCCAAGCTCGAATTAACTCCCTCACTAAAGG-3'

BACKWARD: 5'-CCAGTGAATGTAAACGACTCACTATAGGCG-3'

Seq primer: 5'-GAAATTAACCTCACTAAAGG-3'.

FEATURES

SOURCE

1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Myeloma (MYE) cDNA library"
/sex="male"
/tissue_type="Blood"
/cell_type="myeloma"
/dev_stage="multiple myeloma"

/note="Vector: Lambda Zap Express; Site_1: EcoRI; Site_2: XhoI; Myeloma cells from multiple myeloma patients' bone marrow were purified by magnetic cell sorting. mRNA were purified and an oligo d(T)18 primer containing XhoI restriction site was used to prime first strand synthesis using M-MuV reverse transcriptase. To protect the cDNAs from XhoI digestion in subsequent cloning step, the nucleotide analogue 5-methyl-dCTP was added to the nucleotide mixture and [a-32P]dATP was added to monitor the quantity and quality of first strand synthesis. After second-strand synthesis and blunting of cDNA termini, EcoRI adapters were ligated, followed by kinase treatment and digestion with XhoI. The cDNAs were then size-fractionated using Sephacryl S-500 column and then ligated into EcoRI and XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10⁶. Clones from the primary library were randomly selected for single pass sequencing."

BASE COUNT 34 a 27 c 25 g 24 t

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

alignment_block:

BASK-853-CLAIM4 x A2882861/rev ..

Align seg 1/1 to reverse of: A2882861 from: 1 to: 112

1 GluAlaGly*****Ser 6

BASK-853-CLAIM4 x BF175147 ..

Align seg 1/1 to: BF175147 from: 1 to: 110

1 GluAlaGly*****Ser 6

61 GAGGAGGAGTCATCATCC 78

seq_name: gb_gss:A2882861

seq_documentation_block:

LOCUS A2882861 112 bp DNA linear GSS 05-MAR-2001
DEFINITION RPCI-23-186P2-TV RPCI-23 Mus musculus genomic clone RPCI-23-186P2,
DNA sequence.

ACCESSION A2882861

VERSION A2882861.1 GI:13201806

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 112)

AUTHORS Zhao S., Niernan W., Feldblyum T., Malek J., Shatsman S., Akninet

B., Levins M., McGann S., Tsegaye G., Geer K., Krol M., de Jong, P.

and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end

page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 186 row: P column: 2

Seq primer: T7

Class: BAC ends.

FEATURES

SOURCE

1..112
/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-186P2"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:

EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcoRI and EcoRI Methylase. Size

selected DNA was cloned into the pBACe3.6 vector at the

DH10B sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 20 a 34 c 25 g 33 t

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x A2882861/rev ..

Align seg 1/1 to reverse of: A2882861 from: 1 to: 112

1 GluAlaGly*****Ser 6

34 a 27 c 25 g 24 t

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

```

|||||
67 GAGGCGGGGCGCAAGT 50
seq_name: gb_gss:BH406737

seq_documentation_block:
LOCUS      BH406737
DEFINITION 1007002B12.x1 1007 - RescueMu Grid H Zea mays genomic, DNA
ACCESSION  AL467236.1
VERSION     AL467236.1
KEYWORDS    GI:11837140
SOURCE      Trypanosoma brucei.
ORGANISM    Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1 (bases 1 to 112)
AUTHORS    Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
            Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
            Melville, S.E., Rajandream, M.A., and Barrell, B.G.
TITLE      Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
JOURNAL    project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh@sanger.ac.uk
COMMENT    Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/projects/T_brucei/.

FEATURES             source
            1..112
                /organism="Trypanosoma brucei"
                /strain="TREU927"
                /db_xref="taxon:5691"
                /clone="157b06"
BASE COUNT    36 a 20 c 26 g 30 t
ORIGIN
alignment_scores:
    Quality: 19.00      Length: 6
    Ratio: 3.167       Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x TA157B06P ..
Align seg 1/1 to: TA157B06P from: 1 to: 112

1 GluAlaGly*****Ser 6
|||||
73 GAGGCTGGCGCTGCTCG 90

seq_name: gb_est1:AF090252

seq_documentation_block:
LOCUS      AF090252
DEFINITION AF090252 Rat limonene-treated regressing tumor cDNA library Rattus
            norvegicus cDNA clone mrg-38, mRNA sequence.
ACCESSION  AF090252
VERSION     AF090252
KEYWORDS    GI:3659554
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 114)
AUTHORS    Ariazi, E.A. and Gould, M.N.
TITLE      Identifying differential gene expression in monoterpene-treated
|||||
65 GAGGCGGGGCGCAAGT 48
seq_name: gb_gss:TA157B06P

```

mammary carcinomas using subtractive display
J. Biol. Chem. 271 (46), 29286-29294 (1996)
Contact: Ariazi, E.A. and Gould, M.N.
Oncology
University of Wisconsin-Madison
600 Highland Ave., Room K4/3, Madison, WI 53792, USA
repressed by monoterpene.

FEATURES

source
1. .114
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Wistar-Furth"
/db_xref="taxon:10116"
/clone="mrg-38"
/clone_lib="Rat limonene-treated regressing tumor cDNA library"
/tissue_type="7,12-dimethylbenz[a]anthracene (DMBA)-induced mammary carcinoma"
/note="limonene-treated regressing tumor cDNA library subtracted from a control-treated non-regressing cDNA library"

BASE COUNT 31 a 24 c 32 g 27 t
ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AF090252 ..

Align seg 1/1 to: AF090252 from: 1 to: 114

1 GluAlaGly*****Ser 6
|||||||
96 GAGCGGGCACCTCTCA 113

seq_name: gb_est1:AV624809

seq_documentation_block:

LOCUS AV624809 115 bp mRNA linear EST 15-DEC-2000
DEFINITION AV624809 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas reinhardtii cDNA clone Jc083a09_r 5', mRNA sequence.

ACCESSION AV624809
VERSION AV624809.1 GI:10773986
KEYWORDS EST.

SOURCE

Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE

1 (bases 1 to 115)
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohshima, K., Nakamura, Y., and Tabata, S.

AUTHORS

Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii

TITLE

DNA Res. 7 (5), 303-307 (2000)

JOURNAL

20539644

MEDLINE

Contact: Erika Asamizu

COMMENT

The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source
1. .115
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="JC083a09_r"
/note="Vector: BluescriptII SK-; Site.1: EcoRI; Site.2: XhoI; The cDNA library was constructed from cells cultured

in a carbon stress acclimatized condition in which carbon dioxide concentration in the bubbling gas was changed from 5% to 0.04%

BASE COUNT 23 a 58 c 12 g 22 t
ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AV624809/rev ..

Align seg 1/1 to reverse of: AV624809 from: 1 to: 115

1 GluAlaGly*****Ser 6
|||||||
75 GAGCGAGGATCGTCGCC 58

seq_name: gb_est1:BE170537

seq_documentation_block:

LOCUS BE170537 117 bp mRNA linear EST 21-JUN-2000
DEFINITION QV4-HT0538-170300-142-e12 HT0538 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE170537
VERSION BE170537.1 GI:8633258
KEYWORDS EST.

SOURCE

human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 117)
Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.K., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st5-QV4-HT0538-170300-142-e12st3-2000-03-17st4-1)
Seq primer: puc 18 forward
High quality sequence stop: 117.

FEATURES

source

1. .117
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0538"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

24 a 37 c 32 g 24 t
ORIGIN

```
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BE170537 rev ..
Align seg 1/1 to reverse of: BE170537 from: 1 to: 117

1 GluAlaGly*****Ser 6
|||||
81 GAAGCAGGAGCGCTTCC 64

seq_name: gb_gss:BH096312

seq_documentation_block:
LOCUS BH096312
DEFINITION RPCI-24-391J2-TV RPCI-24 Mus musculus genomic clone RPCI-24-391J2,
DNA sequence.
ACCESSION BH096312
VERSION BH096312.1 GI:14917658
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 117)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html
Plate: 391 row: J column: 2
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..117
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-391J2"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 15 a 25 c 37 g 40 t
ORIGIN
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BH096312 ..
Align seg 1/1 to: BH096312 from: 1 to: 117
```

```
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BE170537 ..
Align seg 1/1 to: BE170537 from: 1 to: 117

1 GluAlaGly*****Ser 6
|||||
58 GAAGCAGGAGCGCTTCC 75

seq_name: gb_est1:BE170537

seq_documentation_block:
LOCUS BE170537
DEFINITION QV4-HT0538-170300-142-e12 HT0538 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE170537
VERSION BE170537.1 GI:8633258
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 117)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-QV4-HT0538-170300-142-e12&t3=2000-03-17&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 117.
Location/Qualifiers
1..117
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0538"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 24 a 37 c 32 g 24 t
ORIGIN
alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167       Gaps: 0
```

```

1 GluAlaGly*****Ser 6
|||||
17 GAGCAGGGGCCAGCTCT 34

seq_name: gb_gss:AZ442930

seq_documentation_block:
LOCUS      118 bp      DNA      linear      GSS 04-OCT-2000
DEFINITION  LM0237D24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0237D24 F, DNA sequence.
ACCESSION  AZ442930
VERSION    AZ442930
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 118)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0237 row: D column: 24
            Seq primer: CGTTGTAAAGACGCGCAGT
            Class: plasmid ends
            High quality sequence stop: 118.
            Location/Qualifiers
FEATURES             1..118
     source           1..118
     /organism="Mus musculus"
     /strain="C57Bl/6J"
     /db_xref="taxon:10090"
     /clone="UUGC1M0237D24"
     /clone_lib="Mouse 10kb plasmid UUGC1M library"
     /sex="Male"
     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
     /notes="Vector: PWD42nv; Purified genomic DNA from M.
            musculus C57Bl/6J (male) was obtained from the Jackson
            Laboratory Mouse DNA Resource
            (http://www.jax.org/resources/documents/dnares/). The DNA
            was hydrodynamically sheared by repeated passage through a
            0.005 inch orifice at constant velocity. The sheared DNA
            was blunt end-repaired with T4 DNA polymerase and T4
            polynucleotide kinase. Adaptor oligonucleotides were
            ligated to the blunt ends in high molar excess. The
            adapted DNA was purified and size-selected for a 9.5 to
            10.5 kb range using preparative agarose gel
            electrophoresis. Vector DNA was prepared from a derivative
            of pWD42 (g14732114|gb|AF129072.1), a copy-number
            inducible derivative of plasmid R1. The vector was ligated
            with adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            chemically-competent E. coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
BASE COUNT      29 a      34 c      19 g      36 t
ORIGIN

alignment_scores:
Quality: 19.00      Length: 6
Ratio: 3.167      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BM434428/rev ..
Align seg 1/1 to reverse of: BM434428 from: 1 to: 119

1 GluAlaGly*****Ser 6
|||||
46 GAGCAGGGTCTGCTCG 29

seq_name: gb_gss:AZ256123

seq_documentation_block:
LOCUS      119 bp      DNA      linear      GSS 26-JUL-2000
DEFINITION  RPCI-23-100B21.TV RPCI-23 Mus musculus genomic clone RPCI-23-100B21

```

```

Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x AZ442930/rev ..
Align seg 1/1 to reverse of: AZ442930 from: 1 to: 118

1 GluAlaGly*****Ser 6
|||||
43 GAAGCTGGCTCCAGTAGC 26

seq_name: gb_est2:BM434428

seq_documentation_block:
LOCUS      119 bp      mRNA      linear      EST 31-JAN-2002
DEFINITION  1RT10E12 Bos taurus Reticulum #1 library Bos taurus cDNA, mRNA
sequence.
ACCESSION  BM434428
VERSION    BM434428.1 GI:18456150
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE  1 (bases 1 to 119)
AUTHORS   Hansen,C., Fu,A., Meng,Y., Li,C., Okine,E., Sensen,C.W., Gordon
            P.M.K. and Moore,S.S.
TITLE      Gene Expression Profiling of the Bovine Gastrointestinal Tract
            Unpublished (2002)
JOURNAL    JOURNAL
COMMENT    Contact: Dr. Stephen Moore
            Beef Genomics Laboratory
            Dept of AFNS, University of Alberta
            410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
            Tel: 780 492 0169
            Fax: 780 492 4265
            Email: smoores@afns.ualberta.ca
            Insert Length: 119 Std Error: 0.00
            POLYA=No.
            Location/Qualifiers
FEATURES             1..119
     source           1..119
     /organism="Bos taurus"
     /db_xref="taxon:9913"
     /clone_lib="Bos taurus Reticulum #1 library"
     /tissue_type="Smooth muscle"
     /cell_type="Stratified squamous epithelial"
     /dev_stage="Young adult"
     /lab_host="XL1-BlueMRF'strain"
     /note="Organ: Reticulum; Vector: Uni-2ZAPXR; Site_1: EcoRI
            ; Site_2: Xho I"
BASE COUNT      17 a      34 c      33 g      35 t
ORIGIN

alignment_scores:
Quality: 19.00      Length: 6
Ratio: 3.167      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BM434428/rev ..
Align seg 1/1 to reverse of: BM434428 from: 1 to: 119

1 GluAlaGly*****Ser 6
|||||
46 GAGCAGGGTCTGCTCG 29

seq_name: gb_gss:AZ256123

seq_documentation_block:
LOCUS      119 bp      DNA      linear      GSS 26-JUL-2000
DEFINITION  RPCI-23-100B21.TV RPCI-23 Mus musculus genomic clone RPCI-23-100B21

```


TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cygaps-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10110 row: g column: 19
 High quality sequence stop: 122.
 Location/Qualifiers

FEATURES
 source
 1..122
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:440218"
 /clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 17 a 44 c 43 g 18 t
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 BASK-853-CLAIM4 x BG032062 ..
 Align seg 1/1 to: BG032062 from: 1 to: 122

1 GluAlaGly*****Ser 6
 ||||||||||||||||
 96 GAAGCGCGCGCAGTCG 113

seq_name: gb_est1:AW880397

seq_documentation_block:
 LOCUS AW880397 125 bp mRNA linear EST 23-MAY-2000
 DEFINITION QV0-OT0030-100400-188-b07 OT0030 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW880397
 VERSION AW880397.1 GI:8042407
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 125)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922

Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=QV0-OT0030-100400-188-b07&t3=2000-04-10&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 12
 High quality sequence stop: 125.
 Location/Qualifiers

FEATURES
 source
 1..125
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="OT0030"
 /dev_stage="Adult"
 /note="Organ: ovary; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 18 a 46 c 29 g 32 t
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 6
 Ratio: 3.167 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 BASK-853-CLAIM4 x AW880397/rev ..
 Align seg 1/1 to reverse of: AW880397 from: 1 to: 125

1 GluAlaGly*****Ser 6
 ||||||||||||||||
 42 GAAGCAGGAGCAGCAGT 25

seq_name: gb_est2:BG383537

seq_documentation_block:
 LOCUS BG383537 125 bp mRNA linear EST 12-MAR-2001
 DEFINITION 301674 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BG383537
 VERSION BG383537.1 GI:13308009
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 125)
 AUTHORS Fahrénkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
 Unpublished (2000)
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACACGTATGACCAT
 BACKWARD: GTTTTCCAGTCACGACG
 Plate: 87 row: J column: 17
 Seq primer: ATTTAGTGACACTATAG.

FEATURES
sourceLocation/Qualifiers
1. .125

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pcMV SPOT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT 24 a 34 c 44 g 23 t
ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BG383537 ..

Align seg 1/1 to: BG383537 from: 1 to: 125

1 GluAlaGly*****Ser 6

|||||
70 GAAGCAGCACCAGCAGC 87

seq_name: gb_est2:BI779080

seq_documentation_block:

LOCUS BI779080 125 bp mRNA linear EST 26-SEP-2001
DEFINITION Ebro01_SQ002_N24_R IGF Barley Ebro01 library Hordeum vulgare cDNA
clone Ebro01_SQ002_N24 5', mRNA sequence.

ACCESSION

BI779080

VERSION

BI779080.1 GI:15781972

KEYWORDS

EST.

SOURCE

barley.

ORGANISM

Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticaceae; Hordeum.

REFERENCE

1 (bases 1 to 125)

Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,

Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.

Development of Barley Transcriptome Resources

Unpublished (2001)

Contact: Waugh R

Unit of Genomics

Scottish Crop Research Institute

Invergowrie, Dundee, DD2 5DA, Scotland, UK

Tel: 00 44 1382 562731

Fax: 00 44 1382 562426

Email: rwaugh@scri.sari.ac.uk

All sequence has a Phred quality score of 20 or over

Seq primer: M13 reverse.

Location/Qualifiers

1. .125

/organism="Hordeum vulgare"

/cultivar="Optic"

/db_xref="taxon:4513"

/clone_lib="IGF Barley Ebro01 library"

/tissue_type="Root, unstressed"

/lab_host="DH10B"

/note="Vector: pSPOT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPOT1.
Derived from roots of 3 week old hydroponically grown
unstressed barley plants. Developed as part of the barley
transcriptome resources of BBSRC/SEERAD funded cereal IGF
(Investigating Gene Function) project."

BASE COUNT 34 a 41 c 33 g 17 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BI779080 ..

Align seg 1/1 to: BI779080 from: 1 to: 125

1 GluAlaGly*****Ser 6

|||||
97 GAGGCGGTCGACGAGT 114

seq_name: gb_est1:AW997381

seq_documentation_block:

LOCUS AW997381 128 bp mRNA linear EST 05-JUN-2000
DEFINITION RCO-BN0050-160200-011-d05 BN0050 Homo sapiens cDNA, mRNA sequence.

ACCESSION

AW997381

VERSION

AW997381.1 GI:8257615

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 128)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RC0-BN0050-160

200-011-d05&t3=2000-02-16&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 12

High quality sequence stop: 128.

FEATURES

Location/Qualifiers

1. .128

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="BN0050"

/dev_stage="Adult"

/note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 13 a 41 c 29 g 45 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

```

Ratio: 3.167          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM4 x AW997381/rev ..
Align seg 1/1 to reverse of: AW997381 from: 1 to: 128

1 GluAlaGly*****Ser 6
|||||
23 GAAGCAGGAGCAGCTCC 6

seq_name: gb_est2:BF983162

seq_documentation_block:
LOCUS   BF983162          128 bp  mRNA  linear  EST 23-JAN-2001
DEFINITION   602309067F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4400249 5',
            mRNA sequence.
ACCESSION   BF983162
VERSION     BF983162.1 GI:12385902
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 128)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLMI0105 row: e column: 18
            High quality sequence stop: 128.
            Location/Qualifiers
                source
                1..128
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4400249"
                /clone_lib="NIH_MGC_88"
                /tissue_type="duodenal adenocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: small intestine; Vector: pCMV-SPORT6;
                Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                oligo-dT primed. Average insert size 1.767 kb. Library
                enriched for full-length clones and constructed by Life
                Technologies. Note: this is a NIH_MGC Library."
                18 a 46 c 46 g 18 t

BASE COUNT      18 a 46 c 46 g 18 t
ORIGIN
1..128
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4400249"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
18 a 46 c 46 g 18 t

alignment_scores:
Quality: 19.00          Length: 6
Ratio: 3.167           Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BF983162 ..
Align seg 1/1 to: BF983162 from: 1 to: 128

1 GluAlaGly*****Ser 6
|||||
102 GAAGCCGGCCAGTTCG 119

seq_name: gb_est2:BG253687

seq_documentation_block:
LOCUS   BG253687          129 bp  mRNA  linear  EST 05-FEB-2002
DEFINITION   602366742F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4474714 5',
            mRNA sequence.
ACCESSION   BG253687
VERSION     BG253687.1 GI:12763503
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 129)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLMI0299 row: d column: 11
            High quality sequence stop: 129.
            Location/Qualifiers
                source
                1..129
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4474714"
                /clone_lib="NIH_MGC_91"
                /tissue_type="adenocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.4 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
                18 a 46 c 46 g 19 t

BASE COUNT      18 a 46 c 46 g 19 t
ORIGIN
1..129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4474714"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
18 a 46 c 46 g 19 t

alignment_scores:
Quality: 19.00          Length: 6
Ratio: 3.167           Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BG253687 ..
Align seg 1/1 to: BG253687 from: 1 to: 129

1 GluAlaGly*****Ser 6
|||||
103 GAAGCCGGCCAGTTCG 120

seq_name: gb_est2:BM445566

seq_documentation_block:
LOCUS   BM445566          129 bp  mRNA  linear  EST 05-FEB-2002
DEFINITION   IL13A7.ab1 Bos taurus ileum #1 library Bos taurus cDNA, mRNA
            sequence.
ACCESSION   BM445566
VERSION     BM445566.1 GI:18529722
KEYWORDS    EST.
SOURCE      cow.
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 129)
AUTHORS    Hansen,C., Fu,A., Meng,Y., Li,C., Okine,E., Sensen,C.W., Gordon

```

TITLE
JOURNAL
COMMENT

,P.M.K. and Moore,S.S.
Gene Expression Profiling of the Bovine Gastrointestinal Tract
Unpublished (2002)
Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: smoores@afns.ualberta.ca
Insert Length: 129 Std Error: 0.00
POLYA-No.

FEATURES
source

1. .129
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="Bos taurus Ileum #1 library"
/tissue_type="Smooth muscle"
/cell_type="Simple columnar epithelial"
/dev_stage="Young adult"
/lab_host="XU1-BlueMRF strain"
/note="Organ: Intestine/Ileum; Vector: Uni-22APXR; Site_1:
ECORI; Site_2: Xho I"
BASE COUNT 28 a 30 c 36 g 35 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BM445566/rev ..

Align seg 1/1 to reverse of: BM445566 from: 1 to: 129

1 GluAlaGly*****Ser 6
|||||
86 GAAGCTGGCAGCAGT 69

seq_name: gb_est2:BG166289

seq documentation_block:
LOCUS BG166289 131 bp mRNA linear EST 06-FEB-2001
DEFINITION 602345480F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4455786 5',
mRNA sequence.
ACCESSION BG166289
VERSION BG166289.1 GI:12672992
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 131)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10249 row: o column: 19
High quality sequence stop: 131.
Location/Qualifiers
1. .131
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

1. .131
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:4455786"
/clone_lib="NIH_MGC_89"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 18 a 47 c 47 g 19 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM4 x BG166289 ..

Align seg 1/1 to: BG166289 from: 1 to: 131

1 GluAlaGly*****Ser 6
|||||
105 GAAGCGGGCCAGTTCG 122

seq_name: gb_est2:BG258468

seq documentation_block:
LOCUS BG258468 131 bp mRNA linear EST 13-FEB-2001
DEFINITION 602379947F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4510666 5',
mRNA sequence.
ACCESSION BG258468
VERSION BG258468.1 GI:12768197
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 131)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10392 row: n column: 11
High quality sequence stop: 131.
Location/Qualifiers
1. .131
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4510666"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

1. .131
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4510666"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 18 a 47 c 47 g 19 t
ORIGIN

alignment_scores:

Quality: 19.00 Length: 6
Ratio: 3.167 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BG258468

Align seg 1/1 to: BG258468 from: 1 to: 131

1 GluAlaGly*****Ser 6

105 GAAGCGCGCGCAGTTCG 122

seq_name: gb_est2:BG403748

seq_documentation_block:

LOCUS BG403748 131 bp mRNA linear EST 12-MAR-2001
DEFINITION 602419577F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4526683 5',
mRNA sequence.

ACCESSION BG403748

VERSION BG403748.1 GI:13297196

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 131)

NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AMI0434 row: 1 column: 20

High quality sequence stop: 131.

FEATURES

source

1..131 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4526683"

/clone_lib="NIH_MGC_93"

/tissue_type="transitional cell papilloma, cell line"

/lab_host="DH10B (phage-resistant)"

/notes="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

18 a 47 c 47 g 19 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BG403748

Align seg 1/1 to: BG403748 from: 1 to: 131

1 GluAlaGly*****Ser 6

105 GAAGCGCGCGCAGTTCG 122

seq_name: gb_est1:AA242543

seq_documentation_block:

LOCUS AA242543 133 bp mRNA linear EST 07-MAR-1997
DEFINITION mx12d05.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:679977 5',
similar to gb:U01024 Mus musculus glycineamide ribonucleotide
synthetase (MOUSE);, mRNA sequence.

ACCESSION AA242543

VERSION AA242543.1 GI:1873327

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 133)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennan,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:419681

Trace considered overall poor quality

Seq primer: -28ml3 rev2 Et from Amersham

High quality sequence stop: 1.

FEATURES

source

1..133 Location/Qualifiers

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:679977"

/clone_lib="Soares mouse NML"

/tissue_type="Liver"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCAATCTTTTTTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

constructed and normalized by Bento Soares and M.Fatima

Bonaldo."

40 a 32 c 38 g 23 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AA242543

Align seg 1/1 to: AA242543 from: 1 to: 133

1 GluAlaGly*****Ser 6

22 GAAGCTGGAGCACATCA 39

seq_name: gb_gss:AZ909098

seq_documentation_block:

AZ909098

LOCUS

133 bp DNA linear GSS 05-MAR-2001

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-NT0201-181200-005-a05&t3=2000-12-18&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 134.

Location/Qualifiers

1. .134

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="MT0201"

/dev_stage="Adult"

/note="Organ: marrow; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

23 a 43 c 34 g 34 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BF903294 ..

Align seg 1/1 to: BF903294 from: 1 to: 134

1 GluAlaGly*****Ser 6

|||||

1 GAGGCGAGTTCGCTCA 18

seq_name: gb_est2:BI856341

seq_documentation_block:

LOCUS BI856341

DEFINITION 603387591F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5396350 5',

mRNA sequence.

ACCESSION BI856341

VERSION BI856341.1 GI:15997088

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 134)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-femail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12010 row: e column: 23

High quality sequence stop: 134.

FEATURES

source

1. .134

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:5396350"

/clone_lib="NIH_MGC_87"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="PH10B (phage-resistant)"

/note="Organ: breast; Vector: pCMV-SPORT6; Site.1: NotI;

Site.2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.383 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

19 a 48 c 47 g 20 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00 Length: 6

Ratio: 3.167 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x BI856341 ..

Align seg 1/1 to: BI856341 from: 1 to: 134

1 GluAlaGly*****Ser 6

|||||

108 GAGCGCGCGCGCTCG 125

seq_name: gb_est2:BG388955

seq_documentation_block:

LOCUS BG388955

DEFINITION 602414709F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4523019 5',

mRNA sequence.

ACCESSION BG388955

VERSION BG388955.1 GI:13282401

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 135)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-femail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10425 row: a column: 04

High quality sequence stop: 135.

Location/Qualifiers

1. .135

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4523019"

FEATURES

source

1. .135

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4523019"

```

/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT      19 a      48 c      48 g      20 t
ORIGIN

```

```

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM4 x BG388955 ..

```

```

Align seg 1/1 to: BG388955 from: 1 to: 135

```

```

1 GluAlaGly*****Ser 6
|||||
109 GAAGCGCGCGCAGTTCG 126

```

```

seq_name: gb_est2:BI090037

```

```

seq_documentation_block:
LOCUS      BI090037      135 bp      mRNA      linear      EST 20-JUN-2001
DEFINITION 602857337F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4998633 5',
mRNA sequence.
ACCESSION  BI090037
VERSION     BI090037.1 GI:14508367
KEYWORDS   EST.
SOURCE     human.

```

```

ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE  1 (bases 1 to 135)
AUTHORS   NIH-MGC http://mgc.nhl.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs.rem@nih.gov
          Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLML1027 row: j column: 10
High quality sequence stop: 135.

```

```

FEATURES   source
            Location/Qualifiers
             1..135
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:4998633"
              /clone_lib="NIH_MGC_10"
              /cell_line="MGC36"
              /lab_host="DH10B"
              /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 1.5 kb. Library prepared by Life
              Technologies."

```

```

BASE COUNT      20 a      49 c      46 g      20 t
ORIGIN

```

```

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0

```

```

Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM4 x AW840788/rev ..

```

```

Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM4 x BI090037 ..

```

```

Align seg 1/1 to: BI090037 from: 1 to: 135

```

```

1 GluAlaGly*****Ser 6
|||||
109 GAAGCGCGCGCAGTTCG 126

```

```

seq_name: gb_est1:AW840788

```

```

seq_documentation_block:
LOCUS      AW840788      136 bp      mRNA      linear      EST 18-MAY-2000
DEFINITION RC4-CN0007-090300-013-b12 CN0007 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW840788
VERSION     AW840788.1 GI:7934771
KEYWORDS   EST.
SOURCE     human.

```

```

ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE  1 (bases 1 to 136)
AUTHORS   Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
          Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
          Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
          Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
          M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
          Simpson, A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=RC4-CN0007-090
          300-013-b12et3=2000-03-09et4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 22
          High quality sequence stop: 136.

```

```

FEATURES   source
            Location/Qualifiers
             1..136
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="CN0007"
              /dev_stage="Adult"
              /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the pUC 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."

```

```

BASE COUNT      26 a      38 c      25 g      47 t
ORIGIN

```

```

alignment_scores:
  Quality: 19.00      Length: 6
  Ratio: 3.167        Gaps: 0

```

```

Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM4 x AW840788/rev ..

```


Align seg 1/1 to reverse of: AW840788 from: 1 to: 136

1 GluAlaGly*****Ser 6
|||||
106 GAAGCTGCTACATCTTCA 89

seq_name: gb_est1:AW922200

seq_documentation_block:

LOCUS AW922200 136 bp mRNA linear EST 19-JUL-2000
DEFINITION LG1_346_H08_b1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA

sequence:

ACCESSION AW922200

VERSION AW922200.1 GI:8088025

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade: Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 136)

Cordonnier-Pratt,M.-M.; Gingle,A.; Marsala,C. and Pratt,L.H.

REFERENCE An EST database from Sorghum: light-grown seedlings

AUTHORS Unpublished (2000)

JOURNAL Contact: Cordonnier-Pratt MM

COMMENT Department of Botany

The University of Georgia

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 542 1805

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16.

The threshold for highest quality sequence

is 20.

Seq primer: JEN REV

High quality sequence start: 2

High quality sequence stop: 132

POLYA-No. Location/Qualifiers

1..136

/organism="Sorghum bicolor"

/db_xref="taxon:4558"

/clone_lib="Light Grown 1 (LGI)"

/notes="Organ: 10- to 14-day-old light-grown (greenhouse)

seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI

; The library was made from poly-A RNA in the cloning

vector lambda Zap II. Clones to be sequenced were

prepared by mass excision."

BASE COUNT 32 a 42 c 35 g 27 t

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.167

Percent Similarity: 100.000

Length: 6

Gaps: 0

Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM4 x AW922200 ..

Align seg 1/1 to: AW922200 from: 1 to: 136

1 GluAlaGly*****Ser 6

|||||

58 GAAGCAGGGGCTGCTCT 75

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:20:56 ; Search time 39.04 Seconds
(without alignments)
5.951 Million cell updates/sec

Title: BASK-853-CLAIM4

Perfect score: 21

Sequence: 1 eagxxs 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	90.5	303	1 HIS1_HAEIN	P43853 haemophilus
2	19	90.5	322	1 YK02 YEAST	P36118 saccharomyc
3	19	90.5	323	1 MAF1 YEAST	P54841 mus musculu
4	19	90.5	323	1 MAF1 RAT	P54842 rattus norv
5	19	90.5	331	1 MACS_HUMAN	P29966 homo sapien
6	19	90.5	352	1 NEPI_ASPO	P46076 aspergillus
7	19	90.5	363	1 AR13_NEUC	Q01356 neurospora
8	19	90.5	367	1 LEU3_BACME	P41019 bacillus me
9	19	90.5	390	1 CARP_SACFI	P22929 saccharomyc
10	19	90.5	443	1 CARA_DICDI	Q9TX43 dictyosteli
11	19	90.5	444	1 VGLX_HSVBS	Q08103 bovine herp
12	19	90.5	474	1 YU88_MYCTU	O53305 mycobacteri
13	19	90.5	482	1 NODT_RHILV	P15727 rhizobium l
14	19	90.5	487	1 ATF2 RAT	Q00969 rattus norv
15	19	90.5	561	1 MERA_ACICA	Q52109 acinetobact
16	19	90.5	561	1 MERA_ENTAG	P94702 enterobacte
17	19	90.5	578	1 AC22_STRCO	P46105 streptomyce
18	19	90.5	636	1 DNK2_SYNY3	P22358 synechocyst
19	19	90.5	638	1 TOXA_PSEAE	P11439 pseudomonas
20	19	90.5	643	1 SGT1_ARATH	Q91sm5 arabidopsis
21	19	90.5	692	1 EOMD_XENLA	P79944 xenopus lae
22	19	90.5	883	1 PGCB_MOUSE	Q61361 mus musculu
23	19	90.5	883	1 PGCB RAT	P55068 rattus norv
24	19	90.5	908	1 SRCA_RABIT	P13666 oryctolagus
25	19	90.5	1012	1 POLS_IBDVA	P08364 avian infec
26	19	90.5	1021	1 TSCC_HUMAN	P50187 homo sapien
27	19	90.5	1027	1 ISWI_DROME	Q24368 drosophila
28	19	90.5	1058	1 PMAL_DICDI	P54679 dictyosteli
29	19	90.5	1069	1 ACNA_ARATH	Q9szr1 arabidopsis
30	19	90.5	1183	1 CNA_STRAU	Q53654 staphylococ
31	19	90.5	1237	1 B3A2_MOUSE	P13808 mus musculu
32	19	90.5	1262	1 MYO6_HUMAN	Q9um54 homo sapien
33	19	90.5	1265	1 DYNA_DROME	P13496 drosophila

34	19	90.5	1265	1 MYO6_MOUSE	Q64331 mus musculu
35	19	90.5	1278	1 Y232_HUMAN	Q26288 homo sapien
36	19	90.5	1395	1 CUT1_MOUSE	P53564 mus musculu
37	19	90.5	1486	1 MUKB_ECOLI	P22523 escherichia
38	19	90.5	1505	1 CUT1_HUMAN	P39880 homo sapien
39	19	90.5	1548	1 UGGG_DROME	Q09332 drosophila
40	19	90.5	2035	1 HMP2_YEREN	P48633 versinia en
41	18	85.7	30	1 CLPA_PNPIS	P81671 pinus pinas
42	18	85.7	99	1 PLAS_CAPBU	P00294 capsella bu
43	18	85.7	123	1 WN14_HUMAN	O14904 homo sapien
44	18	85.7	143	1 PP1B_BACSU	P35137 bacillus su
45	18	85.7	147	1 R157_BOVIN	Q28183 bos taurus
46	18	85.7	154	1 DUT_MYCTU	O07199 mycobacteri
47	18	85.7	155	1 MOAE_RHIME	Q28qz5 rhizobium m
48	18	85.7	156	1 CEST_ECO57	P58233 escherichia
49	18	85.7	157	1 FLIN_PSEAE	Q51466 pseudomonas
50	18	85.7	159	1 YMH2_CAEEL	P34469 caenorhabdi

ALIGNMENTS

RESULT 1

ID	HIS1_HAEIN	STANDARD;	PRT;	303 AA.
AC	P43853;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	ATP phosphoribosyltransferase (EC 2.4.2.17).			
GN	HISG OR HI0468.			
OS	Haemophilus influenzae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;			
OC	Haemophilus.			
ON	NCBI_TaxID=727;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RD / KW20 / ATCC 51907;			
RX	MEDLINE=95350630; PubMed=7542800;			
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,			
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,			
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,			
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,			
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,			
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,			
RA	Fine L.D., Fritchman J.L., Fuhrmann J.L., Geodhagen N.S.M.,			
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,			
RA	Venter J.C.;			
RT	"Whole-genome random sequencing and assembly of Haemophilus			
RT	influenzae Rd."			
RL	Science 269:496-512(1995).			
CC	-!- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribose)-ATP + diphosphate =			
CC	ATP + 5-phospho-alpha-D-ribose 1-diphosphate.			
CC	-!- PATHWAY: FIRST STEP IN HISTIDINE BIOSYNTHETIC PATHWAY. IS VERY			
CC	IMPORTANT IN THE REGULATION OF HISTIDINE METABOLISM.			
CC	-!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-!- SIMILARITY: BELONGS TO THE ATP PHOSPHORIBOSYLTRANSFERASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; U32729; AAC22127.1; -			
DR	TIGR; HI0468; -			
DR	InterPro; IPR001348; HisG.			
DR	Pfam; PF01634; HisG; 1.			
DR	PRODOM; PD003516; HisG; 1.			
DR	PROSITE; PS01316; ATP_F_PHORIBOSYLTR; 1.			

KW Histidine biosynthesis; Transferase; Glycosyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 303 AA; 33821 MW; 08C14D1F6E98A31D CRC64;

Query Match 90.5%; Score 19; DB 1; Length 303;
 Best Local Similarity 66.7%; Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 287 EAGSS 292

RESULT 2

YK02_YEAST ID YK02_YEAST STANDARD; PRT; 322 AA.
 AC P36118;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical 36.6 kDa protein in YPT52-DBP7 intergenic region.
 GN YKR022C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Duesterhoeft A., Moestl D., Poehlmann R., Philippson P.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z28247; CAAB2094.1; -;
 DR EMBL; S38091; S38091.
 DR SGD; S0001730; YKR022C.
 KW Hypothetical protein.
 SQ SEQUENCE 322 AA; 36647 MW; D7A601A46839244C CRC64;

Query Match 90.5%; Score 19; DB 1; Length 322;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 119 EAGSS 124

RESULT 3

MAFL_MOUSE ID MAFL_MOUSE STANDARD; PRT; 323 AA.
 AC P54841;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor MAFL (Segmentation protein KR) (Kreisler).
 GN MAFL OR KRML.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95094266; PubMed=8001130;

RA Cordes S.P., Barsh G.S.;
 RT "The mouse segmentation gene *kr* encodes a novel basic domain-leucine
 RT zipper transcription factor.";
 RL Cell 79:1025-1034(1994).
 CC -1- FUNCTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDBRAIN
 CC SEGMENTATION).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY, GUT, LUNG, AND BRAIN.
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE AT 8.0 DPC (ONE SOMITE) AS A BAND
 CC IN THE CAUDAL HINDBRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES).
 CC THE HIGH LEVEL DOMAIN EXHIBITS A SHARP ROSTRAL EDGE COINCIDENT
 CC WITH THE R4/R5 BOUNDARY AND A DIFFUSE CAUDAL EDGE LOCATED MIDWAY
 CC THROUGH R6.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L36435; AAA65689.1; -;
 DR HSSP; P05412; IJUN.
 DR TRANSFAC; T01439; -;
 DR MGD; MGI:104555; Maf.
 DR InterPro; IPR001871; bZIP.
 DR Pfam; PF03131; bZIP_Maf; 1.
 DR SMART; SM00338; BRLZ; 1.
 DR Transcription regulation; DNA-binding; Nuclear protein.
 FT DOMAIN 131 143
 FT POLY-HIS.
 FT DOMAIN 158 167
 FT POLY-HIS.
 FT DNA_BIND 238 264
 FT BASIC_MOTIF.
 FT DOMAIN 266 287
 FT LEUCINE_ZIPPER.
 FT MUTAGEN 248 248
 FT N->S; LOSS OF TRANSCRIPTIONAL ACTIVITY.
 SQ SEQUENCE 323 AA; 35809 MW; D77AE07ABD9C2AD2 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 323;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 308 EAGSTS 313

RESULT 4

MAFL_RAT ID MAFL_RAT STANDARD; PRT; 323 AA.
 AC P54842;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor MAFL.
 GN MAFL OR MAFL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=Liver;
 RX MEDLINE=97190228; PubMed=9038383;
 RA Sakai M., Imaki J., Yoshida K., Ogata A., Matsushima-Hibaya Y.,
 RA Kuboki Y., Nishizawa M., Nishi S.;
 RT "Rat *maf* related genes: specific expression in chondrocytes, lens and
 RT spinal cord.";
 RL Oncogene 14:745-750(1997).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U56241; AAB50062.1; -
DR HSSP; P05412; LJUN.
DR InterPro; IPR001871; bZIP.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 131 143 POLY-HIS.
FT DOMAIN 158 167 POLY-HIS.
FT DOMAIN 194 201 POLY-ALA.
FT DNA_BIND 238 264 BASIC MOTIF.
FT DOMAIN 266 287 LEUCINE-ZIPPER.
SQ SEQUENCE 323 AA; 35792 MW; 6E386340D1F840A5 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 323;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
DB 308 EAGSTS 313

RESULT 5
MACS_HUMAN STANDARD; PRT; 331 AA.
ID MACS_HUMAN AC P29966;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myristoylated alanine-rich C-kinase substrate (MARCKS) (Protein kinase
DE C substrate, 80 kDa protein, light chain) (PKCSL) (80K-L protein).
GN MACS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91317795; PubMed=1860846;
RA Harlan D.M., Graff J.M., Stumpo D.J., Eddy R.L. Jr., Shows T.B.,
RA Boyle J.M., Blackshear P.J.;
RT "The human myristoylated alanine-rich C kinase substrate (MARCKS)
RT gene (MACS). Analysis of its gene product, promoter, and chromosomal
RT localization.";
RL J. Biol. Chem. 266:14399-14405(1991).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=93052291; PubMed=1427843;
RA Sakai K., Hirai M., Kudoh J., Ninoshima S., Shimizu N.;
RT "Molecular cloning and chromosomal mapping of a cDNA encoding human
RT 80K-L protein; major substrate for protein kinase C.";
RL Genomics 14:175-178(1992).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M68956; AAA59555.1; -
DR EMBL; M68955; AAA59554.1; -
DR EMBL; D10522; BAA01392.1; -
DR PIR; A38873; A38873.
DR MIM; 177061; -
DR InterPro; IPR002101; MARCKS.
DR Pfam; PF02063; MARCKS; 1.
DR PRINTS; PR00963; MARCKS.
DR PROSITE; PS00826; MARCKS.1; 1.
DR PROSITE; PS00827; MARCKS.2; 1.
KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
KW Membrane.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 151 175 CALMODULIN-BINDING (PSD).
FT MOD_RES 158 158 PHOSPHORYLATION (BY PKC).
FT MOD_RES 162 162 PHOSPHORYLATION (BY PKC).
FT MOD_RES 166 166 PHOSPHORYLATION (BY PKC).
FT MOD_RES 169 169 PHOSPHORYLATION (BY PKC).
FT CONFLICT 83 83 S -> A (IN REF. 1).
FT CONFLICT 118 118 A -> P (IN REF. 1).
FT CONFLICT 233 233 P -> S (IN REF. 1).
FT CONFLICT 286 307 PGAPPPQEAAPAEAPAAAAASS ->
FT SEQUENCE 331 AA; 31413 MW; BCC837D586581774 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 331;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
DB 205 EAGAAS 210

RESULT 6
NPIL ASPOR STANDARD; PRT; 352 AA.
ID NPIL ASPOR AC P46076;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neutral protease II precursor (EC 3.4.24.39) (Deuterolysin) (NPIL).
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 176-210; 279-281 AND 304-341.
RX STRAIN=ATCC 20386;
RX MEDLINE=91360097; PubMed=1886621;
RA Tatsumi H., Murakami S., Tsuji R.F., Ishida Y., Murakami K.,
RA Masaki A., Kawabe H., Arimura H., Nakano E., Motai H.;
RT "Cloning and expression in yeast of a cDNA clone encoding Aspergillus
RT oryzae neutral protease II, a unique metalloprotease.";
RL Mol. Genet. 228:97-103(1991).
CC -!- FUNCTION: THERMOSTABLE METALLOPROTEASE. SHOWS HIGH ACTIVITIES ON
CC BASIC NUCLEAR SUBSTRATES SUCH AS HISTONE AND PROTAMINE.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage of bonds with
CC hydrophobic residues in p1'; also 3-Asn-1-Gln-4 and 8-Gln-1-Ser-9
CC bonds in insulin B chain.
CC -!- COFACTOR: BINDS 1 ZINC ION.
CC -!- PTM: PROBABLY POSESSES THREE DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M35 (ZINC
CC METALLOPROTEASE); ALSO KNOWN AS THE DEUTEROLYSIN SUBFAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S53810; AAB19701.1; -
DR MEROPS; M35.002; -
DR InterPro: IPR001384; Peptidase_M35.
DR InterPro: IPR000130; Zn_Mtpeptidse.
DR Pfam: PF02102; Peptidase_M35; 1.
DR PRINTS; PR00768; DEUTEROLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Zymogen.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 175
FT CHAIN 176 352
FT METAL 303 303 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 304 304 BY SIMILARITY.
FT METAL 307 307 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 352 AA; 37517 MW; 070C5131335B7F44 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 352;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 240 EAGSTS 245
||| |

RESULT 7
AR13_NEUCR STANDARD; PRT; 363 AA.
ID AR13_NEUCR
AC Q01356;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Amino acid transporter arg-13.
GN ARG-13.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariata; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDA;
RX MEDLINE=96400914; PubMed=8807290;
RA Liu Q., Dunlap J.C.;
RT "Isolation and analysis of the arg-13 gene of Neurospora crassa.";
RL Genetics 143:1163-1174(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Liu Q., Luo X.;
RT "Phenotypic rescue of Saccharomyces cerevisiae arg11 mutant by
RT Neurospora crassa arg-13 cDNA."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR ARGININE BIOSYNTHESIS. MAY PARTICIPATE IN
CC THE EXPORT OF MATRIX-MADE ORNITHINE INTO THE CYTOSOL (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L36378; AAC37500.1; -

DR EMBL; AF279268; AAF87777.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 266 283 POTENTIAL.
FT TRANSMEM 334 353 POTENTIAL.
SQ SEQUENCE 363 AA; 39401 MW; 8B87A937F6D37DC0 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 363;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
Db 14 EAGAAS 19
||| |

RESULT 8
LEU3_BACME STANDARD; PRT; 367 AA.
ID LEU3_BACME
AC F41019;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE (IMDH) (3-IPM-DH).
GN LEUB OR LEUC.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 319;
RX MEDLINE=94288995; PubMed=7764969;
RA Meinhardt F., Busskamp M., Wittchen K.D.;
RT "Cloning and sequencing of the leu C and npr M genes and a putative
RT spc IV gene from Bacillus megaterium DSM319."
RL Appl. Microbiol. Biotechnol. 41:344-351(1994).
CC -!- CATALYTIC ACTIVITY: 3-CARBOXY-2-HYDROXY-4-METHYLPENTANOATE +
CC NAD(+) = 3-CARBOXY-4-METHYL-2-OXOPENTANOATE + NADH (THE PRODUCT
CC DECARBOXYLATES TO 4-METHYL-2-OXOPENTANOATE).
CC -!- PATHWAY: THIRD STEP IN LEUCINE BIOSYNTHESIS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65184; CAA46295.1; -
DR PIR; S38506; S38506.
DR HSSP; P12010; 2AYQ.
DR InterPro: IPR001804; Isodh.
DR Pfam: PF00180; isodh; 1.
DR PROSITE; PS00470; IDH_INDH; 1.
KW Oxidoreductase; Leucine biosynthesis; NAD.
SQ SEQUENCE 367 AA; 39942 MW; DC04D48E0EEAB0DD CRC64;

Query Match 90.5%; Score 19; DB 1; Length 367;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 eagxxx 6
Db 48 EAGSSS 53

RESULT 9
CARP_SACFI STANDARD; PRT; 390 AA.
AC P22929;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acid protease precursor (EC 3.4.23.-).
GN PEP1.
OS Saccharomycopsis fibuligera (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
OX NCBI_TaxID=4944;
RN [1]

SEQUENCE FROM N.A.
RA Hirata D., Fukui S., Yamashita I.;
RT "Nucleotide sequence of the secretable acid protease gene PEPI in the
RT yeast Saccharomycopsis fibuligera."
RL Agric. Biol. Chem. 52:2647-2649(1988).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00313; BAA00215.1; -.
DR PIR: J03334; J0334.
DR HSP: P32329; LYPS.
DR InterPro: IPR001969; Asp.protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 390 ACID PROTEASE.
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 282 282 BY SIMILARITY.
SQ SEQUENCE 390 AA; 41263 MW; 1350BF97116C54796 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 390;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eagxxx 6
Db 266 EAGSSS 271

RESULT 10
CAR4_DICDI STANDARD; PRT; 443 AA.
AC Q9TX43;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cyclic AMP receptor 4.
GN CARD OR CAR4.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=NC-4;
RX MEDLINE=95047357; PubMed=7958880;
RA Louis J.M., Ginsburg G.T., Kimmel A.R.;
RT "The CAMP receptor CAR4 regulates axial patterning and cellular
RT differentiation during late development of Dictyostelium.";
RL Genes Dev. 8:2086-2096(1994).
CC -1- FUNCTION: RECEPTOR FOR CAMP. REGULATES AXIAL PATTERNING AND
CC CELLULAR DIFFERENTIATION DURING LATE DEVELOPMENT. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: INITIALLY EXPRESSED DURING TIP ELONGATION AND
CC CONTINUES TO ACCUMULATE INTO CULMINATION.
CC -1- PTM: CARBOXYL-TERMINAL SER OR THR RESIDUES MAY BE PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY 5 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC GCRDb; GCR_0277; -.
DR DictyDb; DD07777; CARD.
DR InterPro: IPR000848; GPCR_CAMP.
DR InterPro: IPR000832; GPCR_secretin.
DR Pfam; PF00002; 7tm2; 1.
DR PRINTS; PR00247; GPCR_CAMP.
DR PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Multigene family.
FT DOMAIN 1 11
FT TRANSMEM 12 32 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 33 44 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 45 65 2 (POTENTIAL).
FT DOMAIN 66 89 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 90 110 3 (POTENTIAL).
FT DOMAIN 111 119 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 120 140 4 (POTENTIAL).
FT DOMAIN 141 161 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 162 182 5 (POTENTIAL).
FT DOMAIN 183 208 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 209 229 6 (POTENTIAL).
FT DOMAIN 230 263 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 264 284 7 (POTENTIAL).
FT DOMAIN 285 443 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 293 443 ASN-RICH.
FT DOMAIN 293 298 POLY-ASN.
FT DOMAIN 333 340 POLY-GLN.
FT DOMAIN 343 353 POLY-GLN.
FT DOMAIN 412 426 POLY-ASN.
SQ SEQUENCE 443 AA; 51456 MW; CDF3A9DEE5A5BBE2 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 443;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eagxxx 6
Db 72 EAGSTS 77

RESULT 11
VGLX_HSVBS STANDARD; PRT; 444 AA.
AC Q08103;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein GX precursor.
OS Bovine herpesvirus type 1.2 (strain ST).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=45407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167875; PubMed=8122370;
RA Leung-Tack P., Audonnet J.F., Riviere M.;
RT "The complete DNA sequence and the genetic organization of the short

```

```

RT unique region (US) of the bovine herpesvirus type 1 (ST strain).";
RL Virology 199:409-421(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z23068; CAA80603.1; -
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 444 GLYCOPROTEIN GX.
FT TRANSMEM 390 414 POTENTIAL.
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 444 AA; 46708 MW; 0145942AA35B05CB CRC64;

Query Match 90.5%; Score 19; DB 1; Length 444;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
DB 267 EAGSAS 272

RESULT 12
YU88_MYCTU STANDARD; PRT; 474 AA.
ID YU88_MYCTU STANDARD; PRT; 474 AA.
AC O53305;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 50.9 kDa protein RV3088.
GN RV3088 OR MF3173 OR MFV013.09.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL021309; CAA16146.1; -
DR EMBL; AE007134; AAK47509.1; -
DR TIGR; MT3173; -
DR TubercuList; RV3088; -
DR InterPro; IPR004255; UPF0089.
DR Pfam; PF03007; UPF0089; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 474 AA; 50886 MW; 36832D972BE3851A CRC64;

Query Match 90.5%; Score 19; DB 1; Length 474;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
DB 460 EAGTTS 465

RESULT 13
NODT_RHILV STANDARD; PRT; 482 AA.
ID NODT_RHILV STANDARD; PRT; 482 AA.
AC P15727;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nodulation protein T precursor.
GN NODT.
OS Rhizobium leguminosarum (biovar viciae).
OG Plasmid sym PRL1J1.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=387;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8401;
RX MEDLINE=90151607; PubMed=2303029;
RA Economou A., Hamilton W.D.O., Johnston A.W.B., Downie J.A.;
RT "The Rhizobium nodulation gene noda encodes a Ca2(+)-binding protein
RT that is exported without N-terminal cleavage and is homologous to
RT haemolysin and related proteins.";
RL EMBO J. 9:349-354(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90251164; PubMed=2338917;
RA Surin B.P., Watson J.M., Hamilton W.D.O., Economou A., Downie J.A.;
RT "Molecular characterization of the nodulation gene, noda, from two
RT biovars of Rhizobium leguminosarum.";
RL Mol. Microbiol. 4:245-252(1990).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -!- SIMILARITY: BELONGS TO THE FUSA/NODT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17285; CAA35177.1; -
DR PIR; S08384; S08384.
CC PIR; S10133; S10133.

```


DR InterPro: IPR003423; OEP.
 DR Pfam: PF02321; OEP; 2.
 KW PfamId; Nucleation; Lipoprotein; Signal.
 FT CHAIN 1 17 POTENTIAL.
 FT SIGNAL 18 482 MODULATION PROTEIN T.
 FT LIPID 18 18 N-ACYL DIGLYCERIDE (POTENTIAL).
 SQ SEQUENCE 482 AA; 51488 MW; D555B46566E54F82 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 482;
 Best Local Similarity 66.7%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 ||| |
 Db 201 EAGAAS 206

RESULT 14
 ATF2_RAT STANDARD; PRT; 487 AA.
 AC Q00969; Q62870;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cyclic-AMP-dependent transcription factor ATF-2 (Activating
 DE transcription factor 2) (cAMP response element binding protein CRE-
 DE Bp1).
 GN ATF2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
 RA Muramatsu S.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RA MEDLINE=91332085; PubMed=1714459;
 RA Kageyama R., Sasai Y., Nakanishi S.;
 RT "Molecular characterization of transcription factors that bind to the
 RT CAMP responsive region of the substance P precursor gene. CDNA
 RT cloning of a novel C/EBP-related factor.";
 RL J. Biol. Chem. 266:15525-15531(1991).
 CC -!- FUNCTION: A TRANSCRIPTIONAL ACTIVATOR, PROBABLY CONSTITUTIVE,
 CC WHICH BINDS TO THE CAMP RESPONSIVE REGION (CRE) (CONSENSUS:
 CC 5'GTGACGT(A/C)(A/G)-3'); A SEQUENCE PRESENT IN MANY VIRAL AND
 CC CELLULAR PROMOTERS.
 CC -!- SUBUNIT: BINDS DNA AS A DIMER.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics, and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U38938; AAA93263.1; -
 DR EMBL; M65148; AAA42013.1; -
 DR PIR; A39429; A39429.
 DR HSP; P08047; I3P2.
 DR TRANSFAC; T01382; -
 DR InterPro; IPR000822; Znf-C2H2.1
 DR InterPro; IPR001871; bZIP.
 DR Pfam; PF00170; bZIP; 1.

DR Pfam; PF00096; zf-C2H2; 1.
 DR SMART; SM00338; BRLZ; 1.
 DR SMART; SM00355; Znf-C2H2; 1.
 DR PROSITE; PS00036; BZIP_BASIC; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
 KW Transcription regulation; DNA-binding; Activator; Phosphorylation;
 KW Nuclear protein; Alternative splicing; Zinc-finger; Metal-binding.
 FT ZN_FING 7 31 C2H2-TYPE.
 FT DNA_BIND 333 356 BASIC MOTIF.
 FT DOMAIN 362 390 LEUCINE-ZIPPER.
 FT MOD_RES 51 51 PHOSPHORYLATION (BY MAPK14) (BY
 FT SIMILARITY).
 FT MOD_RES 53 53 PHOSPHORYLATION (BY MAPK14) (BY
 FT SIMILARITY).
 FT VARSPLIC 132 229 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 487 AA; 52286 MW; 4ED95B106DF5F9EE CRC64;

Query Match 90.5%; Score 19; DB 1; Length 487;
 Best Local Similarity 66.7%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 ||| |
 Db 452 EAGATS 457

RESULT 15
 MERA_ACICA STANDARD; PRT; 561 AA.
 AC Q52109;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mercuric reductase (EC 1.16.1.1) (Hg(II) reductase).
 GN MERA.
 OS Acinetobacter calcoaceticus.
 OG Plasmid PKLH2.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OC NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94134837; PubMed=8302940;
 RA Kholodil G.Y., Lomovskaya O.L., Gorlenko Z.M., Mindlin S.Z.,
 RA Yurleva O.V., Nikiforov V.G.;
 RT "Molecular characterization of an aberrant mercury resistance
 RT transposable element from an environmental Acinetobacter strain.";
 RL Plasmid 30:303-308(1993).
 CC -!- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
 CC BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURIC REDUCTASE. MERA
 CC PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
 CC -!- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) = Hg(2+) + NADPH.
 CC -!- COFACTOR: FAD.
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -!- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
 CC -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
 CC OXIDOREDUCTASES CLASS-I.
 CC -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF213017; AAA19682.1; -
 DR HSSP; Q04656; IAW0.
 DR InterPro; IPR001327; FAD_pyr_redox.
 DR InterPro; IPR001934; HMA.

```

DR InterPro; IPR000815; Hg_reductase.
DR InterPro; IPR001100; pyr_redox.
DR InterPro; IPR004099; pyr_redox_dlm.
DR Pfam; PF00403; HMA; 1.
DR Pfam; PF00070; pyr_redox; 1.
DR Pfam; PF02852; pyr_redox_dlm; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00945; HGRDRTASE.
DR PRINTS; PR00411; PNRDRTASEI.
DR PROSITE; PS01047; HMA_1; 1.
DR PROSITE; PS00846; HMA_2; 1.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW Mercuric resistance; Oxidoreductase; Flavoprotein; FAD; NADP;
KW Mercury; Redox-active center; Metal-binding; Plasmid.
FT DOMAIN 1 66 HMA.
FT NP_BIND 100 130 FAD (ADP PART) (PROBABLE).
FT DISULFID 136 141 REDOX-ACTIVE.
FT NP_BIND 393 403 FAD (FLAVIN PART) (BY SIMILARITY).
FT METAL 558 558 HG(2+) (POTENTIAL).
FT METAL 559 559 HG(2+) (POTENTIAL).
SQ SEQUENCE 561 AA; 58558 MW; 111E02A702C157D6 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 561;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |
DB 45 EAGTSS 50

RESULT 16
MERA_ENTAG
ID MERA_ENTAG STANDARD; PRT; 561 AA.
AC P94702;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mercuric reductase (EC 1.16.1.1) (Hg(II) reductase).
GN MERA.
OS Enterobacter agglomerans (Pantoea agglomerans).
OC Bacterioidetes; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OC NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97303088; PubMed=9159519;
RA Yuriyeva O., Kholodil G., Minakhin L., Gorlenko Z., Kalyaeva E.,
RA Mindlin S., Nikiforov V.;
RT "Intercontinental spread of promiscuous mercury-resistance
transposons in environmental bacteria.";
RL Mol. Microbiol. 24:321-329(1997).
CC -!- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
CC BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURIC REDUCTASE. MERA
CC PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
CC -!- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) = Hg(2+) + NADPH.
CC -!- COFACTOR: FAD.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-I.
CC -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M64683; AAA26690.1; -.
 DR EMBL; AL593842; CAC44196.1; -.
 DR InterPro: IPR003662; sub.transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Antibiotic resistance; Transport; Transmembrane.
 FT TRANSMEM 78 98 POTENTIAL.
 FT TRANSMEM 109 129 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 170 190 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 232 252 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 306 326 POTENTIAL.
 FT TRANSMEM 341 361 POTENTIAL.
 FT TRANSMEM 369 389 POTENTIAL.
 FT TRANSMEM 444 464 POTENTIAL.
 FT TRANSMEM 546 566 POTENTIAL.
 SQ SEQUENCE 578 AA; 59772 MW; 56C1DC75E6038B92 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 578;
 Best Local Similarity 66.7%; Pred. No. 4.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 ||| |
 Db 430 EAGTAS 435

RESULT 18

DNK2_SYNY3
 ID DNK2_SYNY3 STANDARD; PRT; 636 AA.
 AC P22358;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Chapterone protein dnk2 (Heat shock protein 70-2) (Heat shock 70 kDa
 DE protein 2) (HSP70-2).
 GN DNK2 OR DNAK OR SL0170.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91093185; PubMed=1670771;
 RA Chitnis P.R., Nelson N.;
 RT "Molecular cloning of the genes encoding two chaperone proteins of
 RT the cyanobacterium Synechocystis sp. PCC 6803.";
 RL J. Biol. Chem. 266:58-65(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugita M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64 to 92% of the genome.";
 RL DNA Res. 2:153-166(1995).
 CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK.
 CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M57518; AAA27287.1; -.
 DR EMBL; D63999; BAA10059.1; -.
 DR PIR; C39025; C39025.
 DR HSP; P04475; 1DG4.
 DR InterPro: IPR001023; HSP70.
 DR Pfam: PF00012; HSP70; 1.
 DR PRINTS; PR00301; HEATSHOCK70.
 DR PROSITE; PS00297; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KW Chapterone; ATP-binding; Heat shock; Multigene family;
 KW Complete proteome.
 SQ SEQUENCE 636 AA; 67614 MW; 33AE4CEBA28F40A CRC64;

Query Match 90.5%; Score 19; DB 1; Length 636;
 Best Local Similarity 66.7%; Pred. No. 5.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
 ||| |
 Db 615 EAGTSS 620

RESULT 19

TOXA_PSEAE
 ID TOXA_PSEAE STANDARD; PRT; 638 AA.
 AC P11439; Q91417;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exotoxin A precursor (NAD-dependent ADP-ribosyltransferase
 DE (EC 2.4.2.-)).
 GN ETA OR PA1148.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-53.
 RX MEDLINE=84194063; PubMed=6201861;
 RA Gray G.L., Smith D.H., Baldrige J.S., Harkins R.N., Vasil M.L.,
 RA Chen E.Y., Heyneker H.L.;
 RT "Cloning, nucleotide sequence, and expression in Escherichia coli of
 RT the exotoxin A structural gene of Pseudomonas aeruginosa.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:2645-2649(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RN Nature 406:959-964(2000).
 RN [3]
 RP ACTIVE SITE.
 RX MEDLINE=87250491; PubMed=2885323;
 RA Carroll S.F., Collier R.J.;
 RT "Active site of Pseudomonas aeruginosa exotoxin A. Glutamic acid 553
 RT is photolabeled by NAD and shows functional homology with glutamic
 RT acid 148 of diphtheria toxin.";
 RL J. Biol. Chem. 262:8707-8711(1987).
 RN [4]
 RP DOMAINS.
 RX MEDLINE=90375493; PubMed=2118903;
 RA Chaudhary V.K., Jinno Y., Galo M.G., Fitzgerald D., Pastan I.;
 RT "Mutagenesis of Pseudomonas exotoxin in identification of sequences

RT responsible for the animal toxicity.";
RL J. Biol. Chem. 265:16306-16310(1990).
RN [5]
RP DOMAINS.
RX MEDLINE=91006124; PubMed=2170123;
RA Bourdenet S., Vacheron M.-J., Guinand M., Michel G., Arminjon F.;
RT "Biochemical and immunochemical studies of proteolytic fragments of
RL exotoxin A from Pseudomonas aeruginosa.";
RN Eur. J. Biochem. 192:379-385(1990).
RX X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 424-638.
RA MEDLINE=96016159; PubMed=7568123;
RL Li M., Dyda F., Benhar I., Pastan I., Davies D.R.;
RT "The crystal structure of Pseudomonas aeruginosa exotoxin domain III
RT with nicotinamide and AMP: conformational differences with the intact
RL exotoxin.";
RN Proc. Natl. Acad. Sci. U.S.A. 92:9308-9312(1995).
RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 424-638.
RA MEDLINE=9629346; PubMed=8692916;
RL Li M., Dyda F., Benhar I., Pastan I., Davies D.R.;
RT "Crystal structure of the catalytic domain of Pseudomonas exotoxin A
RT complexed with a nicotinamide adenine dinucleotide analog:
RL implications for the activation process and for ADP ribosylation.";
RN Proc. Natl. Acad. Sci. U.S.A. 93:6902-6906(1996).
RX -!- FUNCTION: THIS TOXIN IS AN NAD-DEPENDENT ADP-RIBOSYLTRANSFERASE.
CC IT CATALYZES THE TRANSFER OF THE ADP RIBOSYL MOIETY OF OXIDIZED
CC NAD ONTO ELONGATION FACTOR 2 (EF-2) THUS ARRESTING PROTEIN
CC SYNTHESIS.
CC -!- PPM: THE 8 CYSTEINES PARTICIPATE IN INTRACHAIN DISULFIDE BONDS.
CC -!- SIMILARITY: REGIONAL SEQUENCE SIMILARITY AT THE ACTIVE SITE
CC WITH DIPHTHERIA TOXIN (DT).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01397; AAB59097.1; -.
DR EMBL; AE004544; AAG04537.1; -.
DR PIR; A30347; A30347.
DR PDB; 1AER; 10-JUN-96.
DR PDB; 1DMA; 15-SEP-95.
KW Toxin; Signal; Transferase; Glycosyltransferase; NAD; 3D-structure;
KW Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 638
FT DOMAIN 26 277
FT DOMAIN 278 389
FT DOMAIN 390 429
FT DOMAIN 430 638
FT ACT_SITE 465 465
FT ACT_SITE 578 578
FT CONFLICT 2 4
FT CONFLICT 22 22
FT CONFLICT 204 204
FT CONFLICT 389 389
FT CONFLICT 432 432
FT CONFLICT 540 540
FT SEQUENCE 638 AA; 69284 MW; 7B9AAD56A27C700A CRC64;
Query Match 90.5%; Score 19; DB 1; Length 638;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
||| |

Db 384 EGAAS 389
RESULT 20
SGTL_ARATH STANDARD; PRT; 643 AA.
AC Q9LSM5;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE SGT1 protein homolog At5g65490.
GN AT5G65490 OR K1904.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SGT1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB026638; BAA98169.1; -.
DR KW Hypothetical protein.
DR SGT1 SEQUENCE 643 AA; 73161 MW; 9F9DD990F65B4C4F CRC64;
Query Match 90.5%; Score 19; DB 1; Length 643;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 eagxxs 6
||| |
Db 445 EAGSS 450
RESULT 21
EOMD_XENLA STANDARD; PRT; 692 AA.
AC P79944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Eomesodermin.
GN EOMES.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gastrula;
RX MEDLINE=97133207; PubMed=8978604;
RA Ryan K., Garrett N., Mitchell A., Gurdon J.B.;
RT "Eomesodermin, a key early gene in Xenopus mesoderm differentiation.";
RL Cell 87:989-1000(1996).
CC -!- FUNCTION: INVOLVED IN MESODERM DIFFERENTIATION. ACTIVATES WNT8,
CC BRACHYURY, CHD AND MIX.1 EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT OR JUST AFTER MIDBLASTULA

CC TRANSITION (STAGE 8). MAXIMALLY EXPRESSED AT STAGE 10 AS AN
 CC EPIGONIAL MESODERM BAND. MORE PROMINENTLY ON THE DORSAL SIDE
 CC AND AROUND THE INVAGINATING DORSAL LIP.
 CC -1- DOMAIN: BY ACTIVIN.
 CC -1- DOMAIN: CONTAINS 13 S-P-X-X REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U75996; AAC60061.1; -
 CC HSP; P24781; 1XBR.
 CC InterPro: IPR001699; T-box.
 CC Pfam: PF00907; T-box; 1.
 CC PRINTS: PR00937; TBOX.
 CC SMART: SM00425; TBOX; 1.
 CC PROSITE: PS01283; TBOX_1; 1.
 CC PROSITE: PS01264; TBOX_2; 1.
 CC PROSITE: PS0252; TBOX_3; 1.
 CC KW Developmental protein; Transcription regulation; DNA-binding;
 KW Nuclear protein; Repeat.
 FT DNA_BIND 263 443 T-BOX.
 FT SEQUENCE 692 AA; 75943 MW; 9D129A67F6357989 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 692;
 Best Local Similarity 66.7%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
 ||| |
 Db 90 EAGSSS 95

RESULT 22
 PCGB_MOUSE STANDARD; PRT; 883 AA.
 AC Q61361;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Brevican core protein precursor.
 GN BCAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE=97432816; PubMed=92866596;
 RA Rauch U., Meyer H., Brakebusch C., Seidenbecher C., Gundelfinger E.D.,
 RA Beier D.R., Fassler R.;
 RT "Sequence and chromosomal localization of the mouse brevican gene.";
 RL Genomics 44:15-21(1997).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
 CC THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
 CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN.
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
 CC SIMILARITY).
 CC -1- TISSUE SPECIFICITY: BRAIN (BY SIMILARITY).
 CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 IMMUGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X87096; CAA60575.1; -
 CC HSP; P20693; 1HLJ.
 CC MGD; MGI:1096385; Bcan.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003596; Ig_V.
 CC InterPro: IPR000538; Link.
 CC InterPro: IPR000436; Sush1_SCR_CCP.
 CC InterPro: IPR001304; Lectin_c.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF00059; lectin_c; 1.
 CC Pfam; PF00084; sush1; 1.
 CC Pfam; PF00193; Xlink; 2.
 CC ProDom; PD000918; Link; 2.
 CC SMART; SM00032; CCP; 1.
 CC SMART; SM00034; CLECT; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00406; IGV; 1.
 CC SMART; SM00445; LINK; 2.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS00290; IG_MHC; 1.
 CC PROSITE; PS01241; Link; 2.
 CC PROSITE; PS00615; C-TYPE-LECTIN_1; 1.
 CC PROSITE; PS00041; C-TYPE-LECTIN_2; 1.
 CC KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sush1;
 KW EGF-like domain; Repeat; Immunoglobulin domain.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 883 BREVICAN CORE PROTEIN.
 FT DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.
 FT DOMAIN 173 250 LINK 1.
 FT DOMAIN 271 352 LINK 2.
 FT DOMAIN 622 658 EGF-LIKE.
 FT DOMAIN 658 786 C-TYPE LECTIN.
 FT DOMAIN 787 851 SUSHI.
 FT DISULFID 56 136 BY SIMILARITY.
 FT DISULFID 178 249 BY SIMILARITY.
 FT DISULFID 202 223 BY SIMILARITY.
 FT DISULFID 276 351 BY SIMILARITY.
 FT DISULFID 300 321 BY SIMILARITY.
 FT DISULFID 626 637 BY SIMILARITY.
 FT DISULFID 631 646 BY SIMILARITY.
 FT DISULFID 648 657 BY SIMILARITY.
 FT DISULFID 664 675 BY SIMILARITY.
 FT DISULFID 692 784 BY SIMILARITY.
 FT DISULFID 760 776 BY SIMILARITY.
 FT DISULFID 791 834 BY SIMILARITY.
 FT DISULFID 820 847 BY SIMILARITY.
 FT CARBOHYD 129 129 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 883 AA; 96013 MW; CC2C33C97B453E45 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 883;
 Best Local Similarity 66.7%; Pred. No. 7.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
 ||| |
 Db 558 EAGSSS 563

RESULT 23

PCGB_RAT
ID PCGB_RAT STANDARD; PRT; 883 AA.
AC P5068; Q63040; Q62860; Q63513;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brevican core protein precursor (Brain enriched hyaluronan binding
DE protein) (BEHAV protein).
GN BCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=96070828; PubMed=7592978;
RA Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C.,
RA Gundelfinger E.D.;
RT "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as
RT secreted and cell surface glycosylphosphatidylinositol-anchored
RT isoforms.";
RL J. Biol. Chem. 270:27206-27212(1995).
[2]
SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.
RC TISSUE=Brain;
RX MEDLINE=96074575; PubMed=7488217;
RA Yamada H., Watanabe K., Shimomura M., Yamasaki M., Yamauchi Y.;
RT "cDNA cloning and the identification of an aggreganase-like cleavage
RT site in rat brevican.";
RL Biochem. Biophys. Res. Commun. 216:957-963(1995).
[3]
SEQUENCE OF 1-423 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=94216386; PubMed=7512973;
RA Jaworski D.M., Kelly G.M., Hockfield S.;
RT "BEHAV, a new member of the proteoglycan tandem repeat family of
RT hyaluronan-binding proteins that is restricted to the brain.";
RL J. Cell Biol. 125:495-509(1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN. THE
CC GPI-ANCHORED ISOFORM MAY FUNCTION AS A CHONDROITIN SULFATE-
CC BEARING CELL SURFACE RECEPTOR.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX AND ONE FORM
CC ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SECRETED FORM (SHOWN HERE) AND
CC A GPI-ANCHORED FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- DEVELOPMENTAL STAGE: SOLUBLE FORM INCREASES FROM DAY P4 TO P64.
CC GPI-ANCHORED ISOFORM INCREASES AFTER DAY P8.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 364
CC ONWARD AND IS SMALLER (371 AA) DUE TO A FRAMESHIFT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X79881; CAA56255.1; -;
CC EMBL; X86406; CAA60160.1; -;
CC EMBL; U37142; AAA87847.1; -;

DR EMBL; Z28366; CAA82215.1; ALT_FRAME.
DR HSP; P20693; IHLJ.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; xlink; 2.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS01241; LINK; 2.
DR PROSITE; PS00615; C-TYPE-LECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE-LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain; Alternative splicing;
KW GPI-anchor.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 883
FT DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 173 250 LINK 1.
FT DOMAIN 271 352 LINK 2.
FT DOMAIN 622 658 EGF-LIKE.
FT DOMAIN 658 786 C-TYPE LECTIN.
FT DOMAIN 787 851 SUSHI.
FT DISULFID 56 136 BY SIMILARITY.
FT DISULFID 178 249 BY SIMILARITY.
FT DISULFID 202 223 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 300 321 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 657 BY SIMILARITY.
FT DISULFID 791 834 BY SIMILARITY.
FT DISULFID 820 847 BY SIMILARITY.
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 625 645 DCIPSPCHNGTCLKEEGFR -> NSAEQMPAFLFLLL
FT VARSPLIC 646 883 MISSING (IN GPI-ANCHORED ISOFORM).
FT CONFLICT 51 52 AL -> WV (IN REF. 3).
FT CONFLICT 503 503 TV -> L (IN REF. 2).
FT CONFLICT 518 519 TV -> PA (IN REF. 2).
FT CONFLICT 526 526 G -> R (IN REF. 2).
FT CONFLICT 541 541 G -> A (IN REF. 2).
FT CONFLICT 556 556 R -> S (IN REF. 2).
FT CONFLICT 573 573 E -> A (IN REF. 2).
FT CONFLICT 583 583 V -> L (IN REF. 2).
FT CONFLICT 649 649 V -> L (IN REF. 2).
FT CONFLICT 670 670 P -> A (IN REF. 2).
FT CONFLICT 738 738 P -> A (IN REF. 2).
FT CONFLICT 809 809 R -> A (IN REF. 2).
SQ SEQUENCE 883 AA; 96057 MW; AC7ACCA40CB53ED37 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 883;
Best Local Similarity 66.7%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
||| |

```

Db 558 EAGSSS 563

RESULT 24
SRCA_RABIT
ID SRCA_RABIT STANDARD; PRT; 908 AA.
AC P13666;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sarcalumenin precursor.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89345602; PubMed=2762314;
RX Leberer E., Charuk J.H.M., Green N.M., MacLennan D.H.;
RT "Molecular cloning and expression of cDNA encoding a luminal calcium
binding glycoprotein from sarcoplasmic reticulum.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6047-6051(1989).
[2]
RN SEQUENCE OF 1-19 AND 458-908 FROM N.A.
RP MEDLINE=89123480; PubMed=2521635;
RX Leberer E., Charuk J.H.M., Clarke D.M., Green N.M., Zubrycka-Gaarn E.,
RA MacLennan D.H.;
RT "Molecular cloning and expression of cDNA encoding the 53,000-dalton
glycoprotein of rabbit skeletal muscle sarcoplasmic reticulum";
RL J. Biol. Chem. 264:3484-3493(1989).
CC -1- FUNCTION: PERHAPS INVOLVED IN THE REGULATION OF CALCIUM TRANSPORT.
CC -1- SUBCELLULAR LOCATION: SARCOPLASMIC RETICULUM LUMEN. ASSOCIATED
CC THROUGH CA(2+) WITH THE MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A CALCIUM-BINDING
CC GLYCOPROTEIN/160 KDA (SHOWN HERE) AND A SECOND GLYCOPROTEIN/53
CC KDA; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M25750; AAA31189.1; -
DR EMBL; J04480; AAA60730.1; -
DR PIR; A33280; A33280.
DR PIR; A33312; A33312.
KW Calcium-binding; Glycoprotein; Signal; Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 908 150 KDA SARCALUMENIN.
FT CHAIN 458 908 53 KDA SARCALUMENIN.
FT DOMAIN 20 457 ACIDIC DOMAIN, PROBABLY BINDS CALCIUM.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 716 716 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPHYC 21 456 MISSING (IN 53 KDA ISOFORM).
FT CONFLICT 474 474 Q -> E (IN REF. 2).
SQ SEQUENCE 908 AA; 97920 MW; A48CAA221AE1418B CRC64;

Query Match 90.5%; Score 19; DB 1; Length 908;
Best Local Similarity 66.7%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 397 EAGAAS 402

RESULT 25
POLIS_IBDVA
ID POLIS_IBDVA STANDARD; PRT; 1012 AA.
AC P08364;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein [Contains: Major structural protein VP2;
DE Nonstructural protein VP4; Minor structural protein VP3].
OS Avian infectious bursal disease virus (strain Australian 002-73)
(OS IBDV).
OC Viruses; dsRNA viruses; Birnaviridae; Avibirnavirus.
OX NCBI_TaxID=10997;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86229073; PubMed=3014441;
RX Hudson P.J., McKern N.M., Power B.E., Azad A.A.;
RT "Genomic structure of the large RNA segment of infectious bursal
disease virus.";
RL Nucleic Acids Res. 14:5001-5012(1986).
[2]
RN SEQUENCE OF 703-1012 FROM N.A.
RP MEDLINE=86220784; PubMed=3011501;
RX Hudson P.J., McKern N.M., Fahey K.J., Azad A.A.;
RT "Predicted sequence of the host-protective immunogen of infectious
bursal disease virus.";
RL FEBS Lett. 201:143-146(1986).
CC -1- FUNCTION: SEGMENT A ENCODES A POLYPROTEIN, THAT IS PROCESSED INTO
CC THE MAJOR STRUCTURAL PROTEINS OF THE VIRION VP2 AND VP3, AND INTO
CC THE PUTATIVE PROTEASE VP4.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X03993; CAA27629.1; ALT_INIT.
DR PIR; A24382; GNXSAU.
DR MEROPS; S50.002; -
DR InterPro; IPR002662; Birna_VP2.
DR InterPro; IPR002663; Birna_VP3.
DR InterPro; IPR002664; Birna_VP4.
DR Pfam; PF01766; Birna_VP2; 1.
DR Pfam; PF01767; Birna_VP3; 1.
DR Pfam; PF01768; Birna_VP4; 1.
KW Polyprotein; Structural protein; Nonstructural protein; Hydrolase;
KW Protease.
FT CHAIN 1 452 MAJOR STRUCTURAL PROTEIN VP2.
FT CHAIN 454 722 NONSTRUCTURAL PROTEIN VP4 (PROTEASE).
FT CHAIN 724 1012 MINOR STRUCTURAL PROTEIN VP3.
SQ SEQUENCE 1012 AA; 109503 MW; D9320A90459DE8F6 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1012;
Best Local Similarity 66.7%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 484 EAGNAAS 489

RESULT 26
TSSC_HUMAN
ID TSSC_HUMAN STANDARD; PRT; 1021 AA.
AC P55017;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thiazide-sensitive sodium-chloride cotransporter (NA-CL symporter).
GN SLC12A3 OR TSC.
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS GS.
 RX MEDLINE=96122035; PubMed=8528245;
 RA Simon D.B., Nelson-Williams C., Bia M.J., Ellison D., Karet F.E.,
 RA Molina A.M., Vaara I., Iwata F., Cushner H.M., Koolen M., Gainza F.J.,
 RA Gitelman H.J., Lifton R.P.;
 RA "Gitelman's variant of Bartter's syndrome, inherited hypokalaemic
 RT alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl
 RT cotransporter.";
 RL Nat. Genet. 12:24-30(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Kidney;
 RC MEDLINE=37001149; PubMed=8812482;
 RA Mastrolanni N., de Fusco M., Zollo M., Arrigo G., Zuffardi O.,
 RA Bettinelli A., Ballabio A., Casari G.;
 RA "Molecular cloning, expression pattern, and chromosomal localization
 RT of the human Na-Cl thiazide-sensitive cotransporter (SLC12A3).";
 RL Genomics 35:486-493(1996).
 CC -1- FUNCTION: ELECTRICALLY SILENT TRANSPORTER SYSTEM WHICH IS A
 CC MEDIATOR OF SODIUM AND CHLORIDE REABSORPTION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: PREDOMINANT IN KIDNEY.
 CC -1- DISEASE: DEFECTS IN SLC12A3 ARE THE CAUSE OF GITELMAN'S SYNDROME
 CC (GS), AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY DIVERSE
 CC ABNORMALITIES IN ELECTROLYTE HOMEOSTASIS INCLUDING HYPOKALAEMIC
 CC METABOLIC ALKALOSIS. GS IS A SUBSET OF BARTTER'S SYNDROME.
 CC -1- SIMILARITY: BELONGS TO THE SLC12A FAMILY OF TRANSPORTERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U44128; AAC50355.1; -;
 DR EMBL; X91220; CAA62613.1; -;
 DR MIM; 600968; -;
 DR MIM; 263800; -;
 DR InterPro; IPR002293; AA_rel_permease_1.
 DR InterPro; IPR002948; NaCl_transporter.
 DR PRINTS; PR01230; NACLTRNSPORT.
 KW Transport; Transmembrane; Glycoprotein; Disease mutation.
 FT DOMAIN 1 135 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 136 156 POTENTIAL.
 FT DOMAIN 157 158 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 159 179 POTENTIAL.
 FT DOMAIN 180 218 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 219 239 POTENTIAL.
 FT DOMAIN 240 261 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 262 282 POTENTIAL.
 FT DOMAIN 283 286 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 287 307 POTENTIAL.
 FT DOMAIN 308 339 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 340 360 POTENTIAL.
 FT DOMAIN 361 377 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 378 398 POTENTIAL.
 FT DOMAIN 399 452 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 453 473 POTENTIAL.
 FT DOMAIN 474 511 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 512 532 POTENTIAL.
 FT DOMAIN 533 534 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 535 555 POTENTIAL.
 FT DOMAIN 556 577 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 578 598 POTENTIAL.
 FT DOMAIN 599 660 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 661 681 POTENTIAL.
 FT DOMAIN 682 1021 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 209 209 R -> W (IN GS).
 FT VARIANT 349 349 P -> L (IN GS).
 FT VARIANT 421 421 /FTid=VAR_007114.
 FT VARIANT 486 486 C -> R (IN GS).
 FT VARIANT 496 496 D -> N (IN GS).
 FT VARIANT 561 561 /FTid=VAR_007116.
 FT VARIANT 588 588 MISSING (IN GS).
 FT VARIANT 630 630 A -> V (IN GS).
 FT VARIANT 655 655 /FTid=VAR_007119.
 FT VARIANT 728 728 G -> V (IN GS).
 FT VARIANT 741 741 /FTid=VAR_007120.
 FT VARIANT 850 850 R -> H (IN GS).
 FT VARIANT 955 955 /FTid=VAR_007121.
 FT CONFLICT 459 460 R -> L (IN GS).
 FT CONFLICT 766 766 /FTid=VAR_007122.
 FT CONFLICT 807 807 A -> T (IN GS).
 FT CONFLICT 1021 AA; 113138 MW; D7ECE53DA6233821 CRC64;
 SQ SEQUENCE 1021 AA; 113138 MW; D7ECE53DA6233821 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1021;
 Best Local Similarity 66.7%; Pred. No. 8.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 eaqxss 6
 |||||
 Db 121 EAGTSS 126

RESULT 27
 ID ISWI_DROME STANDARD; PRT; 1027 AA.
 AC Q24368; Q9V6E8;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Iswi protein (imitation swi protein) (Nucleosome remodeling factor 140
 DE kDa subunit) (NURF-140) (CHRA140 kDa subunit).
 GN ISWI OR CG8625.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISO-1;
 RX MEDLINE=94187693; PubMed=7908117;
 RA Elfving L.K., Deuring R., McCallum C.M., Peterson C.L., Tamkun J.W.;
 RT "Identification and characterization of Drosophila relatives of the
 RL yeast transcriptional activator SNF2/SWI2.";
 RL Mol. Cell. Biol. 14:2225-2234(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananatzides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahiké C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jastali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sriden-Klamis I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of Drosophila melanogaster";
 RT Science 287:2185-2195(2000).
 RL
 CC -1- FUNCTION: COMPONENT OF THE NUCLEOSOME REMODELING FACTOR COMPLEX
 CC (NURF). A PROTEIN COMPLEX THAT FACILITATES THE PERTURBATION OF
 CC CHROMATIN STRUCTURE IN VITRO IN AN ATP-DEPENDENT MANNER. THE
 CC HYDROLYSIS OF ATP DURING THE REMODELING OF CHROMATIN IS LIKELY TO
 CC BE MEDIATED BY ISWI, RELEASING INORGANIC PHOSPHATE. IT IS ALSO A
 CC COMPONENT OF THE ATP-UTILIZING CHROMATIN ASSEMBLY AND REMODELING
 CC FACTOR (ACF) AND OF THE CHROMATIN ACCESSIBILITY COMPLEX (CHAC).
 CC THIS SUBUNIT MAY SERVE AS THE ENERGY-TRANSDUCING COMPONENT OF
 CC CHROMATIN-REMODELING MACHINES.
 CC
 CC -1- SUBUNIT: NURF IS COMPOSED OF FOUR SUBUNITS; A 215 kDa PROTEIN,
 CC IMITATION SWITCH (ISWI), NURF-55, AND NURF-38.
 CC
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
 CC
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY. SNF2L
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L27127; AAA19868.1; -.
 CC DR EMBL: AE003921; AAF58479.1; -.
 CC DR FlyBase; FBgn0011604; ISWI.
 CC DR InterPro; IPR001410; DEAD.
 CC DR InterPro; IPR001650; Helicase_C.
 CC DR InterPro; IPR001005; Myb_DNA_bind.
 CC DR InterPro; IPR000330; SNF2_N.
 CC DR Pfam; PF00271; helicase_C; 1.
 CC DR Pfam; PF00176; SNF2_N; 1.
 CC DR SMART; SM00487; DEXDC; 1.
 CC DR SMART; SM00490; HELIC; 1.
 CC DR SMART; SM00395; SANT; 2.
 CC Nuclear protein; Helicase; ATP-binding.

FT NP_BIND 153 160 ATP (POTENTIAL).
 FT SITE 256 259 DEAH BOX.
 FT DOMAIN 978 981 POLY-LYS.
 FT DOMAIN 1023 1027 POLY-LYS.
 SQ SEQUENCE 1027 AA; 118873 MW; 008FC81AE15E071F CRC64;

 Query Match 90.5%; Score 19; DB 1; Length 1027;
 Best Local Similarity 66.7%; Pred. No. 8.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 eaqxss 6
 ||| |
 Db 658 EAGTSS 663

 RESULT 28
 PMAL_DICDI STANDARD; PRT; 1058 AA.
 ID PMAL_DICDI
 AC P54679;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable plasma membrane ATPase (EC 3.6.3.6) (Proton pump) (PAT2).
 GN PATB.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycotozoa; Dictyostelid; Dictyostellida; Dictyostellum.
 OX NCBI_TaxID=44689;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=AX3;
 RX MEDLINE=98083743; PubMed=9421912;
 RT Coukell M.B., Montakis J., Cameron A.M.;
 RT "The patB gene of Dictyostelium discoideum encodes a P-type H(+)-
 RT ATPase isoform essential for growth and development under acidic
 RT conditions.";
 RL Microbiology 143:3877-3888(1997).
 CC -1- FUNCTION: THE PLASMA MEMBRANE ATPASE IS A HYDROGEN ION PUMP. THE
 CC PROTON GRADIENT IT GENERATES DRIVES THE ACTIVE TRANSPORT OF
 CC NUTRIENTS BY H+ SYMPORT. THE RESULTING EXTERNAL ACIDIFICATION
 CC AND/OR INTERNAL ALKINIZATION MAY MEDIATE GROWTH RESPONSES.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(IN) = ADP + PHOSPHATE +
 CC H(+)(OUT).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIIA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X98286; CAA66931.1; -.
 CC DR Dictydb; DD00061; patB.
 CC DR InterPro; IPR004014; Cation_ATPase.
 CC DR InterPro; IPR001757; E1-E2_ATPase.
 CC DR InterPro; IPR000695; HATPase.
 CC DR InterPro; IPR001454; Hydrolase.
 CC DR Pfam; PF00690; Cation_ATPase_N; 1.
 CC DR Pfam; PF00122; E1-E2_ATPase; 1.
 CC DR Pfam; PF00702; Hydrolase; 1.
 CC DR PRINTS; PR00119; CATATPASE.
 CC DR PRINTS; PR00120; HATPASE.
 CC DR PROSITE; PS00154; ATPASE_E1_E2; 1.
 CC Hydrolyase; Hydrogen ion transport; Transmembrane; Phosphorylation;
 KW Magnesium; ATP-binding.
 FT DOMAIN 1 212 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 213 232 POTENTIAL.
 FT DOMAIN 233 237 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 238 258 POTENTIAL.

FT DOMAIN 259 387 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 388 407 POTENTIAL.

FT DOMAIN 408 417 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 426 447 POTENTIAL.

FT DOMAIN 448 783 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 784 805 POTENTIAL.

FT DOMAIN 806 810 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 811 833 POTENTIAL.

FT DOMAIN 834 849 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 850 870 POTENTIAL.

FT DOMAIN 871 889 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 890 910 POTENTIAL.

FT DOMAIN 911 922 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 923 943 POTENTIAL.

FT DOMAIN 944 967 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 968 988 POTENTIAL.

FT DOMAIN 989 1058 CYTOPLASMIC (POTENTIAL).

FT MOD_RES 480 480 PHOSPHORYLATION (BY SIMILARITY).

FT METAL 728 728 MAGNESIUM (BY SIMILARITY).

FT METAL 732 732 MAGNESIUM (BY SIMILARITY).

FT DOMAIN 44 55 POLY-GLN.

FT DOMAIN 113 116 POLY-SER.

FT DOMAIN 246 249 POLY-LEU.

SQ SEQUENCE 1058 AA; 117373 MW; CB0E5AB9E9B9AF2 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1058;

Best Local Similarity 66.7%; Pred. No. 8.9e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxks 6

DB 110 EAGSSS 115

RESULT 29

ID ACQA_ARATH STANDARD; PRT; 1069 AA.

AC Q9SR1; Q9M0D3;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Potential calcium-transporting ATPase 10, plasma membrane-type

DE (EC 3.6.3.8) (Ca2+-ATPase, isoform 10).

GN ACAL0 OR A74G29900 OR F27B13.140.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20083488; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Voickaert G.,

RA Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terry N.,

RA Harris B., Ansgore W., Brandt P., Grivell L.A., Rieger M.,

RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,

RA Kreis M., Deleney M., Puigdomenech P., Watson M., Schmidtheini T.,

RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,

RA Langham S.-A., McCullagh B., Bilham L., Robben J.,

RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,

RA Braeken M., Woltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,

RA Weizenecker F., Bothé G., Ramsperger U., Hilbert H., Braun M.,

RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,

RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,

RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,

RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,

RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,

RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,

RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,

RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,

RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,

RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,

RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,

RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,

RA Chedfor F., Cooke R., Berger C., Monfort A., Casacuberta E.,

RA Gbbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,

RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,

RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,

RA Frishman D., Haase D., Lencke K., Meves H.-W., Stocker S.,

RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,

RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,

RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,

RA Stoneking T., Kalicki J., Graves T., Harbott G., Edwards J.,

RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,

RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,

RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,

RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,

RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,

RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,

RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,

RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,

RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,

RA Chen E., Marra M., Martienssen R., McCombie W.R.,

RT 'Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.';

RL Nature 402:769-777(1999)..

CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL INTO THE ENDOPLASMIC RETICULUM (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + CA(2+)(CIS) = ADP + PHOSPHATE + CA(2+)(TRANS).

CC -!- ENZYME REGULATION: ACTIVATED BY CALMODULIN (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- DOMAIN: THE N-TERMINUS CONTAINS AN AUTOINHIBITORY CALMODULIN-BINDING DOMAIN, WHICH BINDS CALMODULIN IN A CALCIUM-DEPENDENT FASHION (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2 ATPASES). SUBFAMILY IIB.

CC -!- CAUTION: THE SEQUENCE CAB43665 DIFFERS FROM THAT SHOWN DUE TO WRONG EXON BOUNDARIES PREDICTED FROM THE GENOMIC SEQUENCE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AL161575; CAB79748.1;

DR EMBL; AL050352; CAB43665.1; ALT_SEQ.

DR HSSP; P04191; 1EUL.

DR InterPro: IPR004014; Cation_ATPase.

DR InterPro: IPR001757; E1-E2_ATPase.

DR InterPro: IPR001454; Hydrolase.

DR Pfam; PF00689; Cation_ATPase_C; 1.

DR Pfam; PF00690; Cation_ATPase_N; 1.

DR Pfam; PF00122; E1-E2_ATPase; 1.

DR Pfam; PF00702; Hydrolase; 1.

DR PRINTS; PR00119; CATATPASE.

DR PROSITE; PS00154; ATPASE_E1_E2; 1.

KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;

KW ATP-binding; Metal-binding; Magnesium; Calmodulin-binding;

KW Multigene family.

FT DOMAIN 1 180 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 181 201 POTENTIAL.

FT DOMAIN 202 219 LUMENAL (POTENTIAL).

FT TRANSMEM 220 240 POTENTIAL.

FT DOMAIN 241 369 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 370 389 POTENTIAL.

FT DOMAIN 390 426 LUMENAL (POTENTIAL).

FT TRANSMEM 427 444 POTENTIAL.

```

FT DOMAIN 445 844 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 845 863 POTENTIAL.
FT DOMAIN 864 874 LUMENAL (POTENTIAL).
FT TRANSSEM 875 895 POTENTIAL.
FT DOMAIN 896 915 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 916 938 POTENTIAL.
FT DOMAIN 939 950 LUMENAL (POTENTIAL).
FT TRANSSEM 952 973 POTENTIAL.
FT DOMAIN 974 991 POTENTIAL.
FT TRANSSEM 992 1013 POTENTIAL.
FT DOMAIN 1014 1023 LUMENAL (POTENTIAL).
FT TRANSSEM 1024 1045 POTENTIAL.
FT DOMAIN 1046 1069 POTENTIAL.
FT TRANSSEM 1070 1091 POTENTIAL.
FT DOMAIN 1092 1113 POTENTIAL.
FT TRANSSEM 1114 1135 POTENTIAL.
FT DOMAIN 1136 1157 POTENTIAL.
FT TRANSSEM 1158 1179 POTENTIAL.
FT DOMAIN 1180 1201 POTENTIAL.
FT TRANSSEM 1202 1223 POTENTIAL.
FT DOMAIN 1224 1245 POTENTIAL.
FT TRANSSEM 1246 1267 POTENTIAL.
FT DOMAIN 1268 1289 POTENTIAL.
FT TRANSSEM 1290 1311 POTENTIAL.
FT DOMAIN 1312 1333 POTENTIAL.
FT TRANSSEM 1334 1355 POTENTIAL.
FT DOMAIN 1356 1377 POTENTIAL.
FT TRANSSEM 1378 1399 POTENTIAL.
FT DOMAIN 1400 1421 POTENTIAL.
FT TRANSSEM 1422 1443 POTENTIAL.
FT DOMAIN 1444 1465 POTENTIAL.
FT TRANSSEM 1466 1487 POTENTIAL.
FT DOMAIN 1488 1509 POTENTIAL.
FT TRANSSEM 1510 1531 POTENTIAL.
FT DOMAIN 1532 1553 POTENTIAL.
FT TRANSSEM 1554 1575 POTENTIAL.
FT DOMAIN 1576 1597 POTENTIAL.
FT TRANSSEM 1598 1619 POTENTIAL.
FT DOMAIN 1620 1641 POTENTIAL.
FT TRANSSEM 1642 1663 POTENTIAL.
FT DOMAIN 1664 1685 POTENTIAL.
FT TRANSSEM 1686 1707 POTENTIAL.
FT DOMAIN 1708 1729 POTENTIAL.
FT TRANSSEM 1730 1751 POTENTIAL.
FT DOMAIN 1752 1773 POTENTIAL.
FT TRANSSEM 1774 1795 POTENTIAL.
FT DOMAIN 1796 1817 POTENTIAL.
FT TRANSSEM 1818 1839 POTENTIAL.
FT DOMAIN 1840 1861 POTENTIAL.
FT TRANSSEM 1862 1883 POTENTIAL.
FT DOMAIN 1884 1905 POTENTIAL.
FT TRANSSEM 1906 1927 POTENTIAL.
FT DOMAIN 1928 1949 POTENTIAL.
FT TRANSSEM 1950 1971 POTENTIAL.
FT DOMAIN 1972 1993 POTENTIAL.
FT TRANSSEM 1994 2015 POTENTIAL.
FT DOMAIN 2016 2037 POTENTIAL.
FT TRANSSEM 2038 2059 POTENTIAL.
FT DOMAIN 2060 2081 POTENTIAL.
FT TRANSSEM 2082 2103 POTENTIAL.
FT DOMAIN 2104 2125 POTENTIAL.
FT TRANSSEM 2126 2147 POTENTIAL.
FT DOMAIN 2148 2169 POTENTIAL.
FT TRANSSEM 2170 2191 POTENTIAL.
FT DOMAIN 2192 2213 POTENTIAL.
FT TRANSSEM 2214 2235 POTENTIAL.
FT DOMAIN 2236 2257 POTENTIAL.
FT TRANSSEM 2258 2279 POTENTIAL.
FT DOMAIN 2280 2301 POTENTIAL.
FT TRANSSEM 2302 2323 POTENTIAL.
FT DOMAIN 2324 2345 POTENTIAL.
FT TRANSSEM 2346 2367 POTENTIAL.
FT DOMAIN 2368 2389 POTENTIAL.
FT TRANSSEM 2390 2411 POTENTIAL.
FT DOMAIN 2412 2433 POTENTIAL.
FT TRANSSEM 2434 2455 POTENTIAL.
FT DOMAIN 2456 2477 POTENTIAL.
FT TRANSSEM 2478 2499 POTENTIAL.
FT DOMAIN 2500 2521 POTENTIAL.
FT TRANSSEM 2522 2543 POTENTIAL.
FT DOMAIN 2544 2565 POTENTIAL.
FT TRANSSEM 2566 2587 POTENTIAL.
FT DOMAIN 2588 2609 POTENTIAL.
FT TRANSSEM 2610 2631 POTENTIAL.
FT DOMAIN 2632 2653 POTENTIAL.
FT TRANSSEM 2654 2675 POTENTIAL.
FT DOMAIN 2676 2697 POTENTIAL.
FT TRANSSEM 2698 2719 POTENTIAL.
FT DOMAIN 2720 2741 POTENTIAL.
FT TRANSSEM 2742 2763 POTENTIAL.
FT DOMAIN 2764 2785 POTENTIAL.
FT TRANSSEM 2786 2807 POTENTIAL.
FT DOMAIN 2808 2829 POTENTIAL.
FT TRANSSEM 2830 2851 POTENTIAL.
FT DOMAIN 2852 2873 POTENTIAL.
FT TRANSSEM 2874 2895 POTENTIAL.
FT DOMAIN 2896 2917 POTENTIAL.
FT TRANSSEM 2918 2939 POTENTIAL.
FT DOMAIN 2940 2961 POTENTIAL.
FT TRANSSEM 2962 2983 POTENTIAL.
FT DOMAIN 2984 3005 POTENTIAL.
FT TRANSSEM 3006 3027 POTENTIAL.
FT DOMAIN 3028 3049 POTENTIAL.
FT TRANSSEM 3050 3071 POTENTIAL.
FT DOMAIN 3072 3093 POTENTIAL.
FT TRANSSEM 3094 3115 POTENTIAL.
FT DOMAIN 3116 3137 POTENTIAL.
FT TRANSSEM 3138 3159 POTENTIAL.
FT DOMAIN 3160 3181 POTENTIAL.
FT TRANSSEM 3182 3203 POTENTIAL.
FT DOMAIN 3204 3225 POTENTIAL.
FT TRANSSEM 3226 3247 POTENTIAL.
FT DOMAIN 3248 3269 POTENTIAL.
FT TRANSSEM 3270 3291 POTENTIAL.
FT DOMAIN 3292 3313 POTENTIAL.
FT TRANSSEM 3314 3335 POTENTIAL.
FT DOMAIN 3336 3357 POTENTIAL.
FT TRANSSEM 3358 3379 POTENTIAL.
FT DOMAIN 3380 3401 POTENTIAL.
FT TRANSSEM 3402 3423 POTENTIAL.
FT DOMAIN 3424 3445 POTENTIAL.
FT TRANSSEM 3446 3467 POTENTIAL.
FT DOMAIN 3468 3489 POTENTIAL.
FT TRANSSEM 3490 3511 POTENTIAL.
FT DOMAIN 3512 3533 POTENTIAL.
FT TRANSSEM 3534 3555 POTENTIAL.
FT DOMAIN 3556 3577 POTENTIAL.
FT TRANSSEM 3578 3599 POTENTIAL.
FT DOMAIN 3600 3621 POTENTIAL.
FT TRANSSEM 3622 3643 POTENTIAL.
FT DOMAIN 3644 3665 POTENTIAL.
FT TRANSSEM 3666 3687 POTENTIAL.
FT DOMAIN 3688 3709 POTENTIAL.
FT TRANSSEM 3710 3731 POTENTIAL.
FT DOMAIN 3732 3753 POTENTIAL.
FT TRANSSEM 3754 3775 POTENTIAL.
FT DOMAIN 3776 3797 POTENTIAL.
FT TRANSSEM 3798 3819 POTENTIAL.
FT DOMAIN 3820 3841 POTENTIAL.
FT TRANSSEM 3842 3863 POTENTIAL.
FT DOMAIN 3864 3885 POTENTIAL.
FT TRANSSEM 3886 3907 POTENTIAL.
FT DOMAIN 3908 3929 POTENTIAL.
FT TRANSSEM 3930 3951 POTENTIAL.
FT DOMAIN 3952 3973 POTENTIAL.
FT TRANSSEM 3974 3995 POTENTIAL.
FT DOMAIN 3996 4017 POTENTIAL.
FT TRANSSEM 4018 4039 POTENTIAL.
FT DOMAIN 4040 4061 POTENTIAL.
FT TRANSSEM 4062 4083 POTENTIAL.
FT DOMAIN 4084 4105 POTENTIAL.
FT TRANSSEM 4106 4127 POTENTIAL.
FT DOMAIN 4128 4149 POTENTIAL.
FT TRANSSEM 4150 4171 POTENTIAL.
FT DOMAIN 4172 4193 POTENTIAL.
FT TRANSSEM 4194 4215 POTENTIAL.
FT DOMAIN 4216 4237 POTENTIAL.
FT TRANSSEM 4238 4259 POTENTIAL.
FT DOMAIN 4260 4281 POTENTIAL.
FT TRANSSEM 4282 4303 POTENTIAL.
FT DOMAIN 4304 4325 POTENTIAL.
FT TRANSSEM 4326 4347 POTENTIAL.
FT DOMAIN 4348 4369 POTENTIAL.
FT TRANSSEM 4370 4391 POTENTIAL.
FT DOMAIN 4392 4413 POTENTIAL.
FT TRANSSEM 4414 4435 POTENTIAL.
FT DOMAIN 4436 4457 POTENTIAL.
FT TRANSSEM 4458 4479 POTENTIAL.
FT DOMAIN 4480 4501 POTENTIAL.
FT TRANSSEM 4502 4523 POTENTIAL.
FT DOMAIN 4524 4545 POTENTIAL.
FT TRANSSEM 4546 4567 POTENTIAL.
FT DOMAIN 4568 4589 POTENTIAL.
FT TRANSSEM 4590 4611 POTENTIAL.
FT DOMAIN 4612 4633 POTENTIAL.
FT TRANSSEM 4634 4655 POTENTIAL.
FT DOMAIN 4656 4677 POTENTIAL.
FT TRANSSEM 4678 4699 POTENTIAL.
FT DOMAIN 4700 4721 POTENTIAL.
FT TRANSSEM 4722 4743 POTENTIAL.
FT DOMAIN 4744 4765 POTENTIAL.
FT TRANSSEM 4766 4787 POTENTIAL.
FT DOMAIN 4788 4809 POTENTIAL.
FT TRANSSEM 4810 4831 POTENTIAL.
FT DOMAIN 4832 4853 POTENTIAL.
FT TRANSSEM 4854 4875 POTENTIAL.
FT DOMAIN 4876 4897 POTENTIAL.
FT TRANSSEM 4898 4919 POTENTIAL.
FT DOMAIN 4920 4941 POTENTIAL.
FT TRANSSEM 4942 4963 POTENTIAL.
FT DOMAIN 4964 4985 POTENTIAL.
FT TRANSSEM 4986 5007 POTENTIAL.
FT DOMAIN 5008 5029 POTENTIAL.
FT TRANSSEM 5030 5051 POTENTIAL.
FT DOMAIN 5052 5073 POTENTIAL.
FT TRANSSEM 5074 5095 POTENTIAL.
FT DOMAIN 5096 5117 POTENTIAL.
FT TRANSSEM 5118 5139 POTENTIAL.
FT DOMAIN 5140 5161 POTENTIAL.
FT TRANSSEM 5162 5183 POTENTIAL.
FT DOMAIN 5184 5205 POTENTIAL.
FT TRANSSEM 5206 5227 POTENTIAL.
FT DOMAIN 5228 5249 POTENTIAL.
FT TRANSSEM 5250 5271 POTENTIAL.
FT DOMAIN 5272 5293 POTENTIAL.
FT TRANSSEM 5294 5315 POTENTIAL.
FT DOMAIN 5316 5337 POTENTIAL.
FT TRANSSEM 5338 5359 POTENTIAL.
FT DOMAIN 5360 5381 POTENTIAL.
FT TRANSSEM 5382 5403 POTENTIAL.
FT DOMAIN 5404 5425 POTENTIAL.
FT TRANSSEM 5426 5447 POTENTIAL.
FT DOMAIN 5448 5469 POTENTIAL.
FT TRANSSEM 5470 5491 POTENTIAL.
FT DOMAIN 5492 5513 POTENTIAL.
FT TRANSSEM 5514 5535 POTENTIAL.
FT DOMAIN 5536 5557 POTENTIAL.
FT TRANSSEM 5558 5579 POTENTIAL.
FT DOMAIN 5580 5601 POTENTIAL.
FT TRANSSEM 5602 5623 POTENTIAL.
FT DOMAIN 5624 5645 POTENTIAL.
FT TRANSSEM 5646 5667 POTENTIAL.
FT DOMAIN 5668 5689 POTENTIAL.
FT TRANSSEM 5690 5711 POTENTIAL.
FT DOMAIN 5712 5733 POTENTIAL.
FT TRANSSEM 5734 5755 POTENTIAL.
FT DOMAIN 5756 5777 POTENTIAL.
FT TRANSSEM 5778 5799 POTENTIAL.
FT DOMAIN 5800 5821 POTENTIAL.
FT TRANSSEM 5822 5843 POTENTIAL.
FT DOMAIN 5844 5865 POTENTIAL.
FT TRANSSEM 5866 5887 POTENTIAL.
FT DOMAIN 5888 5909 POTENTIAL.
FT TRANSSEM 5910 5931 POTENTIAL.
FT DOMAIN 5932 5953 POTENTIAL.
FT TRANSSEM 5954 5975 POTENTIAL.
FT DOMAIN 5976 5997 POTENTIAL.
FT TRANSSEM 6000 6021 POTENTIAL.
FT DOMAIN 6022 6043 POTENTIAL.
FT TRANSSEM 6044 6065 POTENTIAL.
FT DOMAIN 6066 6087 POTENTIAL.
FT TRANSSEM 6088 6109 POTENTIAL.
FT DOMAIN 6110 6131 POTENTIAL.
FT TRANSSEM 6132 6153 POTENTIAL.
FT DOMAIN 6154 6175 POTENTIAL.
FT TRANSSEM 6176 6197 POTENTIAL.
FT DOMAIN 6198 6219 POTENTIAL.
FT TRANSSEM 6220 6241 POTENTIAL.
FT DOMAIN 6242 6263 POTENTIAL.
FT TRANSSEM 6264 6285 POTENTIAL.
FT DOMAIN 6286 6307 POTENTIAL.
FT TRANSSEM 6308 6329 POTENTIAL.
FT DOMAIN 6330 6351 POTENTIAL.
FT TRANSSEM 6352 6373 POTENTIAL.
FT DOMAIN 6374 6395 POTENTIAL.
FT TRANSSEM 6396 6417 POTENTIAL.
FT DOMAIN 6418 6439 POTENTIAL.
FT TRANSSEM 6440 6461 POTENTIAL.
FT DOMAIN 6462 6483 POTENTIAL.
FT TRANSSEM 6484 6505 POTENTIAL.
FT DOMAIN 6506 6527 POTENTIAL.
FT TRANSSEM 6528 6549 POTENTIAL.
FT DOMAIN 6550 6571 POTENTIAL.
FT TRANSSEM 6572 6593 POTENTIAL.
FT DOMAIN 6594 6615 POTENTIAL.
FT TRANSSEM 6616 6637 POTENTIAL.
FT DOMAIN 6638 6659 POTENTIAL.
FT TRANSSEM 6660 6681 POTENTIAL.
FT DOMAIN 6682 6703 POTENTIAL.
FT TRANSSEM 6704 6725 POTENTIAL.
FT DOMAIN 6726 6747 POTENTIAL.
FT TRANSSEM 6748 6769 POTENTIAL.
FT DOMAIN 6770 6791 POTENTIAL.
FT TRANSSEM 6792 6813 POTENTIAL.
FT DOMAIN 6814 6835 POTENTIAL.
FT TRANSSEM 6836 6857 POTENTIAL.
FT DOMAIN 6858 6879 POTENTIAL.
FT TRANSSEM 6880 6901 POTENTIAL.
FT DOMAIN 6902 6923 POTENTIAL.
FT TRANSSEM 6924 6945 POTENTIAL.
FT DOMAIN 6946 6967 POTENTIAL.
FT TRANSSEM 6968 6989 POTENTIAL.
FT DOMAIN 6990 7011 POTENTIAL.
FT TRANSSEM 7012 7033 POTENTIAL.
FT DOMAIN 7034 7055 POTENTIAL.
FT TRANSSEM 7056 7077 POTENTIAL.
FT DOMAIN 7078 7099 POTENTIAL.
FT TRANSSEM 7100 7121 POTENTIAL.
FT DOMAIN 7122 7143 POTENTIAL.
FT TRANSSEM 7144 7165 POTENTIAL.
FT DOMAIN 7166 7187 POTENTIAL.
FT TRANSSEM 7188 7209 POTENTIAL.
FT DOMAIN 7210 7231 POTENTIAL.
FT TRANSSEM 7232 7253 POTENTIAL.
FT DOMAIN 7254 7275 POTENTIAL.
FT TRANSSEM 7276 7297 POTENTIAL.
FT DOMAIN 7298 7319 POTENTIAL.
FT TRANSSEM 7320 7341 POTENTIAL.
FT DOMAIN 7342 7363 POTENTIAL.
FT TRANSSEM 7364 7385 POTENTIAL.
FT DOMAIN 7386 7407 POTENTIAL.
FT TRANSSEM 7408 7429 POTENTIAL.
FT DOMAIN 7430 7451 POTENTIAL.
FT TRANSSEM 7452 7473 POTENTIAL.
FT DOMAIN 7474 7495 POTENTIAL.
FT TRANSSEM 7496 7517 POTENTIAL.
FT DOMAIN 7518 7539 POTENTIAL.
FT TRANSSEM 7540 7561 POTENTIAL.
FT DOMAIN 7562 7583 POTENTIAL.
FT TRANSSEM 7584 7605 POTENTIAL.
FT DOMAIN 7606 7627 POTENTIAL.
FT TRANSSEM 7628 7649 POTENTIAL.
FT DOMAIN 7650 7671 POTENTIAL.
FT TRANSSEM 7672 7693 POTENTIAL.
FT DOMAIN 7694 7715 POTENTIAL.
FT TRANSSEM 7716 7737 POTENTIAL.
FT DOMAIN 7738 7759 POTENTIAL.
FT TRANSSEM 7760 7781 POTENTIAL.
FT DOMAIN 7782 7803 POTENTIAL.
FT TRANSSEM 7804 7825 POTENTIAL.
FT DOMAIN 7826 7847 POTENTIAL.
FT TRANSSEM 7848 7869 POTENTIAL.
FT DOMAIN 7870 7891 POTENTIAL.
FT TRANSSEM 7892 7913 POTENTIAL.
FT DOMAIN 7914 7935 POTENTIAL.
FT TRANSSEM 7936 7957 POTENTIAL.
FT DOMAIN 7958 7979 POTENTIAL.
FT TRANSSEM 7980 8001 POTENTIAL.
FT DOMAIN 8002 8023 POTENTIAL.
FT TRANSSEM 8024 8045 POTENTIAL.
FT DOMAIN 8046 8067 POTENTIAL.
FT TRANSSEM 8068 8089 POTENTIAL.
FT DOMAIN 8090 8111 POTENTIAL.
FT TRANSSEM 8112 8133 POTENTIAL.
FT DOMAIN 8134 8155 POTENTIAL.
FT TRANSSEM 8156 8177 POTENTIAL.
FT DOMAIN 8178 8199 POTENTIAL.
FT TRANSSEM 8200 8221 POTENTIAL.
FT DOMAIN 8222 8243 POTENTIAL.
FT TRANSSEM 8244 8265 POTENTIAL.
FT DOMAIN 8266 8287 POTENTIAL.
FT TRANSSEM 8288 8309 POTENTIAL.
FT DOMAIN 8310 8331 POTENTIAL.
FT TRANSSEM 8332 8353 POTENTIAL.
FT DOMAIN 8354 8375 POTENTIAL.
FT TRANSSEM 8376 8397 POTENTIAL.
FT DOMAIN 8398 8419 POTENTIAL.
FT TRANSSEM 8420 8441 POTENTIAL.
FT DOMAIN 8442 8463 POTENTIAL.
FT TRANSSEM 8464 8485 POTENTIAL.
FT DOMAIN 8486 8507 POTENTIAL.
FT TRANSSEM 8508 8529 POTENTIAL.
FT DOMAIN 8530 8551 POTENTIAL.
FT TRANSSEM 8552 8573 POTENTIAL.
FT DOMAIN 8574 8595 POTENTIAL.
FT TRANSSEM 8596 8617 POTENTIAL.
FT DOMAIN 8618 8639 POTENTIAL.
FT TRANSSEM 8640 8661 POTENTIAL.
FT DOMAIN 8662 8683 POTENTIAL.
FT TRANSSEM 8684 8705 POTENTIAL.
FT DOMAIN 8706 8727 POTENTIAL.
FT TRANSSEM 8728 8749 POTENTIAL.
FT DOMAIN 8750 8771 POTENTIAL.
FT TRANSSEM 8772 8793 POTENTIAL.
FT DOMAIN 8794 8815 POTENTIAL.
FT TRANSSEM 8816 8837 POTENTIAL.
FT DOMAIN 8838 8859 POTENTIAL.
FT TRANSSEM 8860 8881 POTENTIAL.
FT DOMAIN 8882 8903 POTENTIAL.
FT TRANSSEM 8904 8925 POTENTIAL.
FT DOMAIN 8926 8947 POTENTIAL.
FT TRANSSEM 8948 8969 POTENTIAL.
FT DOMAIN 8970 8991 POTENTIAL.
FT TRANSSEM 8992 9013 POTENTIAL.
FT DOMAIN 9014 9035 POTENTIAL.
FT TRANSSEM 9036 9057 POTENTIAL.
FT DOMAIN 9058 9079 POTENTIAL.
FT TRANSSEM 9080 9101 POTENTIAL.
FT DOMAIN 9102 9123 POTENTIAL.
FT TRANSSEM 9124 9145 POTENTIAL.
FT DOMAIN 9146 9167 POTENTIAL.
FT TRANSSEM 9168 9189 POTENTIAL.
FT DOMAIN 9190 9211 POTENTIAL.
FT TRANSSEM 9212 9233 POTENTIAL.
FT DOMAIN 9234 9255 POTENTIAL.
FT TRANSSEM 9256 9277 POTENTIAL.
FT DOMAIN 9278 9299 POTENTIAL.
FT TRANSSEM 9300 9321 POTENTIAL.
FT DOMAIN 9322 9343 POTENTIAL.
FT TRANSSEM 9344 9365 POTENTIAL.
FT DOMAIN 9366 9387 POTENTIAL.
FT TRANSSEM 9388 9409 POTENTIAL.
FT DOMAIN 9410 9431 POTENTIAL.
FT TRANSSEM 9432 9453 POTENTIAL.
FT DOMAIN 9454 9475 POTENTIAL.
FT TRANSSEM 9476 9497 POTENTIAL.
FT DOMAIN 9498 9519 POTENTIAL.
FT TRANSSEM 9520 9541 POTENTIAL.
FT DOMAIN 9542 9563 POTENTIAL.
FT TRANSSEM 9564 9585 POTENTIAL.
FT DOMAIN 9586 9607 POTENTIAL.
FT TRANSSEM 9608 9629 POTENTIAL.
FT DOMAIN 9630 9651 POTENTIAL.
FT TRANSSEM 9652 9673 POTENTIAL.
FT DOMAIN 9674 9695 POTENTIAL.
FT TRANSSEM 9696 9717 POTENTIAL.
FT DOMAIN 9718 9739 POTENTIAL.
FT TRANSSEM 9740 9761 POTENTIAL.
FT DOMAIN 9762 9783 POTENTIAL.
FT TRANSSEM 9784 9805 POTENTIAL.
FT DOMAIN 9806 9827 POTENTIAL.
FT TRANSSEM 9828 9849 POTENTIAL.
FT DOMAIN 9850 9871 POTENTIAL.
FT TRANSSEM 9872 9893 POTENTIAL.
FT DOMAIN 9894 9915 POTENTIAL.
FT TRANSSEM 9916 9937 POTENTIAL.
FT DOMAIN 9938 9959 POTENTIAL.
FT TRANSSEM 9960 9981 POTENTIAL.
FT DOMAIN 9982 10000 POTENTIAL.

```



```

DR Pfam: PF00063; myosin_head; 1
DR PRINTS: PR00193; MYOSINHEAVY; 1
DR ProDom: PD000355; myosin_head; 1
DR SMART: SM00015; IQ; 1
DR SMART: SM00242; MYSC; 1
DR PROSITE: PS00096; IQ; FALSE-NEG.
KW Myosin: ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Disease mutation; Deafness.
FT DOMAIN 1 759
FT DOMAIN 814 834
FT DOMAIN 848 1030
FT NP_BIND 151 158
FT DOMAIN 665 672
FT VARIANT 442 442 C->Y (IN DENA22).
FT /FTID=VAR_012110.
SQ SEQUENCE 1262 AA; 146047 MW; CF1FA35796FC1C60 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1262;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
Db 354 EAGSTS 359

RESULT 33
DYNA_DROME
ID DYNA_DROME STANDARD; PRT; 1265 AA.
AC P13496; Q9VUAL;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 150 kDa dynein-associated polypeptide (DP-150) (DAP-150) (Glued
DE protein).
GN GL OR CG9206.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R, AND CANTON-S;
RX MEDLINE=87317680; PubMed=2819881;
RA Swarcop A., Swarcop M., Garen A.;
RT "Sequence analysis of the complete cDNA and encoded polypeptide for
RT the Glued gene of Drosophila melanogaster";
RL Proc. Natl. Acad. Sci. U.S.A. 84:6501-6505(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arif J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.R., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

```

```

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Zhao Q., Zheng L.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: REQUIRED FOR THE CYTOPLASMIC DYNEIN-DRIVEN RETROGRADE
CC MOVEMENT OF VESICLES AND ORGANELLES ALONG MICROTUBULES. DYNEIN-
CC DYNACTIN INTERACTION IS A KEY COMPONENT OF THE MECHANISM OF AXONAL
CC TRANSPORT OF VESICLES AND ORGANELLES.
CC -1- SUBUNIT: LARGE MACROMOLECULAR COMPLEX OF AT LEAST 10 COMPONENTS.
CC P150(GLUED) BINDS DIRECTLY TO MICROTUBULES AND TO CYTOPLASMIC
CC DYNEIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS THE DYNACTIN 150 KDA SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC FRAMESHIFTS AT POSITIONS 32; 174 TO 220; 548 TO 672 AND 1208.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J02932; -; NOT_ANNOTATED_CDS.
CC PIR: AE003536; AAF49788.1; -
CC PIR: A28313; A28313.
CC FlyBase: FBgn0001108; GL.
CC InterPro: IPR000938; CAP-GLY.
CC Pfam: PF01302; CAP_GLY; 1.
CC PROSITE: PS00845; CAP_GLY_1; 1.
CC PROSITE: PS00245; CAP_GLY_2; 1.
KW Motor protein; Microtubules; Dynein; Coiled coil; Cytoskeleton.
FT DOMAIN 27 69
FT DOMAIN 105 138 SER-RICH.
FT DOMAIN 213 570 COILED COIL (POTENTIAL).
FT DOMAIN 812 836 COILED COIL (POTENTIAL).
FT DOMAIN 967 1084 COILED COIL (POTENTIAL).
FT DOMAIN 1128 1160 COILED COIL (POTENTIAL).
FT CONFLICT 708 708 D -> A (IN REF. 1).
FT CONFLICT 875 875 L -> V (IN REF. 1).
FT CONFLICT 888 888 A -> R (IN REF. 1).
FT CONFLICT 1043 1043 S -> C (IN REF. 1).
SQ SEQUENCE 1265 AA; 141217 MW; 2038A200282B2755 CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1265;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
Db 802 EAGATS 807

RESULT 34
MYO6_MOUSE

```

```

ID MYO6_MOUSE STANDARD; PRT; 1265 AA.
AC Q6431;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Myosin VI.
GN MYO6 OR SV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96083582; PubMed=7493015;
RA Avraham K.B., Hasson T., Steel K.P., Kingsley D.M., Russell L.B.,
RT "The mouse Snell's waltzer deafness gene encodes an unconventional
RL myosin required for structural integrity of inner ear hair cells.";
RL Nat. Genet. 11:369-375(1995).
CC -1- FUNCTION: RECESSIVE ACTIN-BASED MOTOR. REQUIRED FOR STRUCTURAL
CC INTEGRITY OF INNER EAR HAIR CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. WITHIN THE COCHLEA,
CC EXPRESSED SPECIFICALLY WITHIN THE SENSORY HAIR CELLS.
CC -1- DISEASE: DEFECTS IN MYO6 ARE THE CAUSE OF SNELL'S WALTZER, A
CC CONDITION CHARACTERIZED BY CIRCLING, HEAD-TOSSING, DEAFNESS AND
CC HYPERACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSTIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49739; AAB00194.1; -.
DR HSSP: P08799; 1LVK.
DR MGD: MGI:104785; Myo6.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR PRINTS: PR00193; MYOSIN_HEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: P550096; IQ; FALSE_NEG.
DR Myosin: ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Disease mutation; Deafness.
FT DOMAIN 1 762 MYOSIN HEAD-LIKE.
FT DOMAIN 817 837 IQ.
FT DOMAIN 849 1014 COILED COIL (POTENTIAL).
FT NP_BIND 151 158 ATP (POTENTIAL).
FT DOMAIN 668 675 ACTIN-BINDING (POTENTIAL).
FT VARIANT 766 1265 MISSING (IN SNELL'S WALTZER).
SQ SEQUENCE 1265 AA; 146408 MW; 4F51ABC72463148C CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1265;
Best Local Similarity 66.7%; Pred. No. 1le+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
DB 355 EAGSTS 360

RESULT 35
Y232_HUMAN STANDARD; PRT; 1278 AA.
ID Y232_HUMAN
RP SEQUENCE FROM N.A. (ISOFORM 1).
RN [1]
RC STRAIN=C57BL/6N;
RA Wang Z., Goldstein A., Neufeld E.J., Scheuermann R.H., Tucker P.W.;
RT "Repression of immunoglobulin heavy chain intronic enhancer
RT through nuclear matrix attachment sites: Cux/CDP homeoprotein is a
RL component of NF-muNR repressor.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).
```

```

AC Q92628;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA0232 (Fragment).
GN KIAA0232.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 271-1278 FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D86985; BAA13221.2; -.
DR Hypothetical protein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 1278 AA; 141663 MW; 2FCFC8837AF8134D CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1278;
Best Local Similarity 66.7%; Pred. No. 1le+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
DB 177 EAGSSS 182

RESULT 36
CUTL_MOUSE STANDARD; PRT; 1395 AA.
AC P53564; O08994; P70301;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CCAAT displacement protein (CDP) (Cut-like 1) (Homeobox protein Cux)
DE (Fragment).
GN CUTL1 OR CUX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6N;
RA Wang Z., Goldstein A., Neufeld E.J., Scheuermann R.H., Tucker P.W.;
RT "Repression of immunoglobulin heavy chain intronic enhancer
RT through nuclear matrix attachment sites: Cux/CDP homeoprotein is a
RL component of NF-muNR repressor.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).
```

RC STRAIN-A/J, AND BALB/C; TISSUE=Brain;
 RX MEDLINE=9424481; PubMed=7910552;
 RA Valarche I., Tissier-Seta J.P., Hirsch M.R., Martinez S., Goridis C.,
 RA Brunet J.F.;
 RT "The mouse homeodomain protein Phox2 regulates Ncam promoter activity
 in concert with Cux/CDP and is a putative determinant of
 neurotransmitter phenotype.";
 RL Development 119:881-896(1993).
 [3]
 RN SEQUENCE OF 642-1395 FROM N.A.
 RP MEDLINE=96437626; PubMed=8840273;
 RA den Heuvel G.B., Bodmer R., McConnell K.R., Nagami G.T., Igarashi P.;
 RT "Expression of a cut-related homeobox gene in developing and
 polycystic mouse kidney.";
 RL Kidney Int. 50:453-461(1996).
 [4]
 RN SEQUENCE OF 936-1395 FROM N.A.
 RP TISSUE=Testis;
 RA Quaglin S.E., Igarashi P.;
 RT "A unique variant of a homeobox gene related to Drosophila cut is
 expressed in mouse testis.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
 REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
 PREVENTING BINDING OF POSITIVELY-ACTING CCAAT FACTORS TO
 PROMOTERS (BY SIMILARITY). COMPONENT OF NF-MUNR REPRESSOR; BINDS
 TO THE MARS (5' AND 3') OF THE IMMUNOGLOBULIN HEAVY CHAIN
 ENHANCER.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 3 CUT DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEOBOX PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF004225; AAD12485.1; -;
 DR EMBL; X75013; CAA52922.1; -;
 DR EMBL; U46683; AAC52775.1; -;
 DR EMBL; U46684; AAB41146.1; -;
 DR HSPSP; P10037; IAU7.
 DR MGD; MGI:88568; Cutl1.
 DR InterPro; IPR003350; CUT.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF02376; CUT; 3.
 DR Pfam; PF00046; homeobox; 2.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Transcription regulation; Homeobox; DNA-binding;
 KW Developmental protein; Nuclear protein; Receptor; Repressor;
 KW Coiled coil; Alternative splicing.
 FT NON_TER 1 1
 FT DOMAIN 1 243 COILED COIL (POTENTIAL).
 FT DNA_BIND 420 507 CUT 1.
 FT DOMAIN 547 603 COILED COIL (POTENTIAL).
 FT DNA_BIND 809 896 CUT 2.
 FT DNA_BIND 992 1079 CUT 3.
 FT VARSPIC 1119 1178 HOMEBOX.
 FT CONFLICT 287 388 MISSING (IN ISOFORM 2).
 FT CONFLICT 1360 1360 G->A (IN REF. 2).
 FT CONFLICT 1365 1365 P->L (IN REF. 1).
 SQ SEQUENCE 1395 AA; 15180 MW; D062CC227D7A163E CRC64;

Query Match 90.5%; Score 19; DB 1; Length 1395;
 Best Local Similarity 66.7%; Pred. No. 1.2e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 eagxxx 6
 Db 377 EAGSTS 382
 III I
 RESULT 37
 MUKB_ECOLI
 ID MUKB_ECOLI STANDARD; PRT; 1486 AA.
 AC P22523; P77164; O47398;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cell division protein mukB.
 GN MUKB OR B0924.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RX MEDLINE=91114703; PubMed=1989883;
 RA Niki H., Jaffe A., Imamura R., Ogura T., Hiraga S.;
 RT "The new gene mukB codes for a 177 kd protein with coiled-coil
 domains involved in chromosome partitioning of E. coli.";
 RL EMBO J. 10:183-193(1991).
 [2]
 RN SEQUENCE FROM N.A. AND MUTANTS MUKB33 AND MUKB106.
 RP MEDLINE=95080615; PubMed=7988894;
 RA Yamanaka K., Mitani T., Feng J., Ogura T., Niki H., Hiraga S.;
 RT "Two mutant alleles of mukB, a gene essential for chromosome
 partition in Escherichia coli.";
 RL FEMS Microbiol. Lett. 123:27-31(1994).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12 / MGI655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 [5]
 RN SEQUENCE OF 1-44 FROM N.A.
 RP STRAIN-K12 / W3110;
 RX MEDLINE=94232180; PubMed=7513784;
 RA Feng J., Yamanaka K., Niki H., Ogura T., Hiraga S.;
 RT "New killing system controlled by two genes located immediately
 upstream of the mukB gene in Escherichia coli.";
 RL Mol. Gen. Genet. 243:136-147(1994).
 [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-227.
 RX MEDLINE=20015369; PubMed=10545328;
 RA van den Ent F., Lockhart A., Kendrick-Jones J., Loewe J.;
 RT "Crystal structure of the N-terminal domain of MukB: a protein
 involved in chromosome partitioning.";
 RL Structure 7:1181-1187(1999).

(UDP-Glc:glycoprotein glucosyltransferase) (UGT).

UGT OR UGGG.

OS Drosophila melanogaster (Fruit fly).

OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Platygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

LN [1]

SEQUENCE FROM N.A., AND SEQUENCE OF 23-37.

TISSUE=Embryo;

MEDLINE=95246722; PubMed=7729408;

RA Parker C.G., Fessler L.I., Nelson R.E., Fessler J.H.;

RT "Drosophila UDP-glucose:glycoprotein glucosyltransferase: sequence

RT and characterization of an enzyme that distinguishes between

RT genaturated and native proteins";

RL EMBO J. 14:1294-1303(1995).

CC -! FUNCTION: UNFOLDED, DENATURED GLYCOPROTEINS ARE SUBSTANTIALLY

CC BETTER SUBSTRATES FOR GLUCOSYLATION BY THIS ENZYME THAN ARE THE

CC CORRESPONDING NATIVE PROTEINS. THIS PROTEIN AND TRANSIT

CC GLYCOSYLATION MAY BE INVOLVED IN MONITORING AND/OR ASSISTING THE

CC FOLDING AND ASSEMBLY OF NEWLY MADE GLYCOPROTEINS, IN ORDER TO

CC IDENTIFY GLYCOPROTEINS THAT NEED ASSISTANCE IN FOLDING FROM

CC CHAPERONES.

CC -! COFACTOR: REQUIRES CALCIUM AND MANGANESE IONS FOR ACTIVITY.

CC -! PATHWAY: GLYCOSYLATION.

CC -! SUBUNIT: MONOMER.

CC -! SUBCELLULAR LOCATION: Endoplasmic reticulum.

CC -! DEVELOPMENTAL STAGE: IS PRESENT AT LOW BUT DETECTABLE LEVELS IN

CC THE EARLIEST EMBRYOS, INCREASING AT 6-8 HRS WITH A MAXIMUM AT 10-

CC 12 HRS. LEVELS DECREASE THEREAFTER AND ARE NOT DETECTED IN 18-20

CC HRS EMBRYOS AND FIRST INSTAR LARVAE BUT IS DETECTED AGAIN AT

CC SECOND INSTAR TO PUPATION!

CC -! SIMILARITY: SOME, TO YEAST KRE5, S.TYPHIMURIUM RFAJ AND E.COLI

CC RFAJ PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U20554; AAA85850.1; -.

CC FlyBase; FBgn0014075; Ugt.

CC InterPro; IPRO02495; Glycosyl_transf_8.

CC Pfam; PF01501; Glyco.transf.8; 1.

CC Signal; Transferase; Glycosyltransferase; Endoplasmic reticulum;

CC Glycoprotein.

CC SIGNAL 1 22

CC CHAIN 23 1548

CC UDP-GLUCOSE:GLYCOPROTEIN

CC GLYCOSYLTRANSFERASE.

CC CARBOHYD 181 181

CC N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 266 266

CC N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 864 864

CC N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SITE 1545 1548

CC PREVENT SECRETION FROM ER (POTENTIAL).

CC SEQUENCE 1548 AA; 174465 MW; 95D6849961622DB6 CRC64;

CC SQ

Query Match	90.5%	Score 19:	DB 1:	Length 1548:
Best Local Similarity	66.7%	Pred. No.	1.3e+03:	
Matches 4: Conservative	0:	Mismatches	2:	Indels 0:
		caps	0:	

Qy 1 eagxs 6
||| |
Db 255 EAGSTS 260

RESULT	40
HMP2_YEREN	
ID	HMP2_YEREN
AC	P48633;
DT	01-FEB-1996 (Rel. 33, Created)
STANDARD;	
PRT;	2035 AA.

01-FEB-1996 (Rel. 33, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE High-molecular-weight protein 2 (HMWP2).
GN IRP2.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OC NCBI_TaxID=630;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=8081 / SEROTYPE O:8;
RX MEDLINE=93374844; PubMed=8366034;
RA Gullvout I., Mercereau-Puljeon O., Bonnefoy S., Pugsley A.P.,
RN Carniel E.;
RT "High-molecular-weight protein 2 of Yersinia enterocolitica is
RT homologous to Angr of Vibrio anguillarum and belongs to a family of
RT proteins involved in nonribosomal peptide synthesis.";
RL J. Bacteriol. 175:5488-5504(1993).
CC -1- FUNCTION: UNKNOWN. MAY BE INVOLVED IN THE NONRIBOSOMAL SYNTHESIS
CC OF SMALL PEPTIDES.
CC -1- COFACTOR: CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC (POTENTIAL).
CC -1- DOMAIN: CONSISTS OF A CENTRAL REGION WITH SIMILARITY TO THE REPEAT
CC DOMAINS OF ACVS AND GRC2, FLANKED BY TWO REPEAT DOMAINS, EACH OF
CC WHICH CONTAINS 5 DIRECT REPEATS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L18881; AAA27636.1; -;
DR EMBL; Z35454; CAA84606.1; -;
DR PIR; A48654; A48654.
DR HSP; P14687; IAMU.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR001242; DUF4.
DR InterPro; IPR003880; Phosphopant_attach.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00668; Condensation; 2.
DR Pfam; PF00550; pp-binding; 3.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
KW Ligase; Multifunctional enzyme; Phosphopantetheine; Repeat.
KW DOMAIN 3 547 I.
FT REPEAT 114 146 I-DR1.
FT REPEAT 310 321 I-DR2.
FT REPEAT 378 390 I-DR3.
FT REPEAT 454 462 I-DR4.
FT REPEAT 477 491 I-DR5.
FT DOMAIN 1466 1919 II.
FT REPEAT 1495 1527 II-DR1.
FT REPEAT 1682 1693 II-DR2.
FT REPEAT 1750 1762 II-DR3.
FT REPEAT 1826 1834 II-DR4.
FT REPEAT 1849 1863 II-DR5.
FT DOMAIN 20 88 ACYL CARRIER (ACP) 1.
FT DOMAIN 1409 1475 ACYL CARRIER (ACP) 2.
FT DOMAIN 1944 2014 ACYL CARRIER (ACP) 3.
FT BINDING 52 52 PHOSPHOPANTETHEINE (BY SIMILARITY).
SQ SEQUENCE 2035 AA; 228826 MW; 1C801377A4375BDC CRC64;

Query Match

90.5%; Score 19; DB 1; Length 2035;

Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 1972 EAGATS 1977

RESULT 41
CLPA_PINPS STANDARD; PRT; 30 AA.
ID CLPA_PINPS
AC P81671;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ATP-dependent c1p protease ATP-binding subunit c1pa homolog
(Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -1- FUNCTION: MAY INTERACT WITH A CLPP-LIKE PROTEASE INVOLVED IN
CC DEGRADATION OF DENATURED PROTEINS IN THE CHLOROPLAST (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC (SPOT N9) IS: 5.9, ITS MW IS: 92 kDa.
CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.
DR InterPro: IPR001270; CLP_AB.
DR PROSITE: PS00870; CLPB_1; PARTIAL.
DR PROSITE: PS00871; CLPB_2; PARTIAL.
KW Chaperone; ATP-binding; Repeat; Chloroplast.
FT NON_TER 1 1
FT NON_CONS 15 16
FT NON_TER 30 30
SQ SEQUENCE 30 AA; 2923 MW; 44B5950B73A96152 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 30;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 4 EAGDAS 9

RESULT 42
PLAS_CAPBU STANDARD; PRT; 99 AA.
ID PLAS_CAPBU
AC P00294;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Plastocyanin.
GN PETE.
OS Capsella bursa-pastoris (Shepherd's purse).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Capsella.
OX NCBI_TaxID=3719;
RN [1]
RP SEQUENCE.
RA Scaven M.D., Ramshaw J.A.M., Brown R.H., Boulter D.;

Unpublished results, cited by:
RL Boulter D., Haslett B.G., Peacock D., Ramshaw J.A.M., Scaven M.D.;
RL (In) Northcote D.H. (eds.);
RL Plant biochemistry II, pp.13:1-40, University Park Press,
RL Baltimore (1977).
CC -1- FUNCTION: PLASTOCYANIN PARTICIPATES IN ELECTRON TRANSFER BETWEEN
CC P700 AND THE CYTOCHROME B/F COMPLEX IN PHOTOSYSTEM I.
CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
CC MEMBRANE SURFACE IN CHLOROPLASTS.
CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
DR PIR: A00304; CUSU.
DR HSP; P00289; 2ECF.
DR Mendel; 11575; CAPBU; PetE; 1.
DR InterPro: IPR001235; Copper_Blue.
DR InterPro: IPR000923; Copper_Blue.
DR Pfam: PF00127; copper-bind; 1.
DR PRINTS: PR00156; COPPERBLUE.
DR PRODOM: PD001235; Copper_Blue; 1.
DR PROSITE: PS00196; COPPER_BLUE; 1.
KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane.
FT DOMAIN 1 99 PLASTOCYANIN-LIKE.
FT METAL 37 37 COPPER (BY SIMILARITY).
FT METAL 84 84 COPPER (BY SIMILARITY).
FT METAL 87 87 COPPER (BY SIMILARITY).
FT METAL 92 92 COPPER (BY SIMILARITY).
SQ SEQUENCE 99 AA; 10383 MW; 30BA97B58B9580F1 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 76 EAGTYS 81

RESULT 43
WN14_HUMAN STANDARD; PRT; 123 AA.
ID WN14_HUMAN
AC O14904;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE WNT-14 protein (Fragment).
GN WNT14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98110581; PubMed=9441749;
RA Bergstein I., Eisenberg L.M., Bhalarao J., Jenkins N.A.,
RA Copeland N.G., Osborne M.P., Bowcock A.M., Brown A.M.C.;
RT "Isolation of two novel WNT genes, WNT14 and WNT15, one of which
RT (WNT15) is closely linked to WNT3 on human chromosome 17q21.";
RL Genomics 46:450-458(1997).
CC -1- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN
CC TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A
CC SIGNALING MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE
CC REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL
CC DIAMETERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC extracellular matrix.
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF028702; AAC39550.1;
DR MIM; 602863;
DR InterPro; IPR000970; Wnt1.
DR Pfam; PF00110; wnt; 1.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; PARTIAL.
KW Developmental protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 13143 MW; 8F000D2568EAE744 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 123;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
Db 43 EAGALS 48

RESULT 44
PIIB_BACSU STANDARD; PRT; 143 AA.
AC P35137;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peptidyl-prolyl cis-trans isomerase B (EC 5.2.1.8) (PPIase B)
DE (Rotanase B).
GN PIIB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / MAREBURG;
RX MEDLINE=95020538; PubMed=7934829;
RA Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;
RT "The organization of the Bacillus subtilis 168 chromosome region
RT between the spoVA and serA genetic loci, based on sequence data.";
RL Mol. Microbiol. 10:385-395(1993).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=168 / JH642;
RX MEDLINE=94293776; PubMed=8022278;
RA Herrler M., Bang H., Marahiel M.A.;
RT "Cloning and characterization of ppIB, a Bacillus subtilis gene which
RT encodes a cyclosporin A-sensitive peptidyl-prolyl cis-trans
RT isomerase.";
RL Mol. Microbiol. 11:1073-1083(1994).
RN [3]
RP SEQUENCE OF 1-26.
RC STRAIN=168 / JH642;
RX MEDLINE=96345629; PubMed=8755892;
RA Graumann P., Schroeder K., Schmid R., Marahiel M.A.;
RT "Cold shock stress-induced proteins in Bacillus subtilis.";
RL J. Bacteriol. 178:4611-4619(1996).
CC -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS.
CC -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -!- ENZYME REGULATION: INHIBITED BY CYCLOSPORIN A (CSA).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CYCLOPHILIN-TYPE PPIASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L09228; AAA67475.1;
DR EMBL; X73898; CAA52103.1;
DR EMBL; Z99116; CAB14268.1;
DR PIR; S45537; S45537.
DR HSSP; Q27450; IA58.
DR Subtilist; BG10512; ppIB.
DR InterPro; IPR002130; CSA_PPIase.
DR Pfam; PF00160; pro_isomerase; 1.
DR PRINTS; PR00153; CSAPPISMRASE.
DR PROSITE; PS00170; CSA_PPIASE_1; 1.
DR PROSITE; PS00072; CSA_PPIASE_2; 1.
KW Isomerase; Rotanase; Complete proteome.
SQ SEQUENCE 143 AA; 15256 MW; 9EF17D70EB81EC51 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 143;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
Db 82 EAGALS 87

RESULT 45
R157_BOVIN STANDARD; PRT; 147 AA.
AC Q28183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Retina-specific 15.7 kDa protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=86259168; PubMed=2425311;
RA Nakagawa Y., Kuo C.H., Ishii K., Shiosaka S., Tohyama M., Miki N.;
RT "Cloning and characterization of a cDNA specific for bovine retina.";
RL Neurosci. Res. 3:300-310(1986).
CC -!- TISSUE SPECIFICITY: RETINA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M34915; AAA30756.1;
SQ SEQUENCE 147 AA; 15658 MW; 1FEDA4878B39645 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 147;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
Db 133 EAGTVS 138

RESULT 46
DUT_MYCTU STANDARD; PRT; 154 AA.
ID DUT_MYCTU

```

```

AC 007199;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)
DE (dUTPase) (dUTP pyrophosphatase).
GN DUT OR RV2697C OR MT2771 OR MTCY05A6.18C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekalia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers R.,
RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: THIS ENZYME IS INVOLVED IN NUCLEOTIDE METABOLISM: IT
CC PRODUCES DUMP, THE IMMEDIATE PRECURSOR OF THYMIDINE NUCLEOTIDES
CC AND IT DECREASES THE INTRACELLULAR CONCENTRATION OF DUTP SO THAT
CC URACIL CANNOT BE INCORPORATED INTO DNA (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: dUTP + H(2)O = dUMP + diphosphate.
CC -1- PATHWAY: DE NOVO SYNTHESIS OF THYMIDYLATE.
CC -1- SIMILARITY: BELONGS TO THE DUTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z96072; CAB09487.1; -.
DR EMBL; AE007106; AAK47086.1; -.
DR HSSP; P06968; 1EUW.
DR TIGR; MT2771; -.
DR TubercuList; RV2697c; -.
DR InterPro; IPR001428; dUTPase.
DR Pfam; PF00692; dUTPase; 1.
DR ProDom; PD000946; dUTPase; 1.
DR Hydrolase; Nucleotide metabolism; Complete proteome.
SQ SEQUENCE 154 AA; 15803 MW; 836D5E6420EF455B CRC64;

Query Match 85.7%; Score 18; DB 1; Length 154;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
DB 132 EAGLAS 137

RESULT 48
CEST_ECO57 ID CEST_ECO57 STANDARD; PRT; 156 AA.
AC P58233;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tlr chaparrone.
GN CEST OR Z5111 OR ECS4560 OR L0026.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;

```

```

RESULT 47
MOAE_RHIME STANDARD; PRT; 155 AA.
AC O92QX5;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Molybdopterin converting factor subunit 2 (MPT synthase subunit 2)
DE (Molybdopterin synthase subunit 2) (Molybdenum cofactor biosynthesis
DE protein E) (Molybdopterin converting factor large subunit).
GN MOAE OR R01168 OR SMC00599.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RA "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- FUNCTION: Converts molybdopterin precursor 2 into molybdopterin.
CC This requires the incorporation of two sulfur atoms into precursor
CC Z to generate ac dithiolene group (by similarity).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SUBUNIT: Heterodimer of 2 moad subunits and 2 moae subunits (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE MOAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL591786; CAC45747.1; -.
DR InterPro; IPR003448; Moae.
DR Pfam; PF02391; Moae; 1.
DR Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 155 AA; 16926 MW; 4DE035E/ADFBB367 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 155;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
DB 43 EAGALS 48

RESULT 48
CEST_ECO57 ID CEST_ECO57 STANDARD; PRT; 156 AA.
AC P58233;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tlr chaparrone.
GN CEST OR Z5111 OR ECS4560 OR L0026.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;

```

RN SEQUENCE FROM N.A.
 RP STRAIN-O157:H7 / HAI;
 RC MEDLINE=96079490; PubMed=8566710;
 RA Zhao S., Mitchell S.E., Meng J., Doyle M.P., Kresovich S.;
 RT "Cloning and nucleotide sequence of a gene upstream of the eaeA gene
 of enterohemorrhagic Escherichia coli O157:H7";
 RL FEMS Microbiol. Lett. 133:35-39(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / ATCC 700927;
 RX MEDLINE=98339885; PubMed=9673266;
 RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
 RA Kaper J.B., Blattner F.R.;
 RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
 Escherichia coli O157:H7";
 RL Infect. Immun. 66:3810-3817(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 RN CC
 CC -1- FUNCTION: Chaperone for the type III secretion of Tir. Probably
 stabilizes the protein, prevents inappropriate protein-
 protein interactions and aids in secretion (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE CEST/SYCH CHAPERONE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32312; AAB00110.1; -;
 CC EMBL: AF071034; AAC31505.1; -;
 CC EMBL: AE005595; AAG58824.1; -;
 CC EMBL: AP002566; BAB37983.1; -;
 CC Chaperone; Virulence.
 KW Chaperone; Virulence.
 SQ SEQUENCE 156 AA; 17681 MW; 999545426E26D2D6 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 156;
 Best Local Similarity 66.7%; Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaaxxs 6
 |||||
 Db 87 EGAQS 92

RESULT 49
 FLIN_PSEAE

ID FLIN_PSEAE STANDARD; PRT; 157 AA.
 AC Q51466;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar motor switch protein flin.
 GN FLIN OR PA1444.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAK;
 RX MEDLINE=95347807; PubMed=7622217;
 RA Simpson D.A., Ramphal R., Lory S.;
 RT "Characterization of pseudomonas aeruginosa flin, a gene involved in
 flagellar biosynthesis and adherence";
 RL Infect. Immun. 63:2950-2957(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Miziouchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-wadman S., Yuan Y.,
 RA Brody L.B., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 opportunistic pathogen";
 RL Nature 406:959-964(2000).
 RN CC
 CC -1- FUNCTION: FLIN IS ONE OF THREE PROTEINS (FLIG, FLIN, FLIM) THAT
 FORM A SWITCH COMPLEX THAT IS PROPOSED TO BE LOCATED AT THE BASE
 OF THE BASIS BODY. THIS COMPLEX INTERACTS WITH THE CHEY AND CHEZ
 CHEMOTAXIS PROTEINS, IN ADDITION TO CONTACTING COMPONENTS OF THE
 MOTOR THAT DETERMINE THE DIRECTION OF FLAGELLAR ROTATION (BY
 SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
 CC -1- SIMILARITY: BELONGS TO THE FLIN/MOPA/SPAO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L39832; AAA79754.1; -;
 CC EMBL: AE004574; AAG04833.1; -;
 CC InterPro: IPR001172; Flag_FLIN.
 CC InterPro: IPR001543; SpOA.
 CC Pfam: PF01052; SpOA; 1.
 CC PRINTS: PR00956; FLGMOTOREFLIN.
 CC ProDom: PD001777; SpOA; 1.
 CC Chemotaxis; Flagella; Flagellar rotation; Inner membrane;
 KW Complete proteome.
 FT CONFLICT 48 48 P -> S (IN REF. 1).
 SQ SEQUENCE 157 AA; 16620 MW; B3D91C0182ACB775 CRC64;

Query Match 85.7%; Score 18; DB 1; Length 157;
 Best Local Similarity 66.7%; Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaaxxs 6
 |||||
 Db 24 EAGDAS 29

RESULT 50
 YMH2_CAEEL

```

ID YMH2_CAEEL STANDARD; PRT; 159 AA.
AC P34469;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 17.2 kDa protein F58A4.2 in chromosome III.
GN F58A4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22179; CA80168.1; -.
DR PIR; S40974; S40974.
DR WormPep; F58A4.2; CE01017.
KW Hypothetical protein.
SQ SEQUENCE 159 AA; 17201 MW; 364FE35A65E2C89D CRC64;

```

```

Query Match      85.7%; Score 18; DB 1; Length 159;
Best Local Similarity 66.7%; Pred. NO. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 eaqx 6
   |||||
Db 145 EAGSGS 150

```

Search completed: September 24, 2002, 11:27:06
Job time: 370 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:14:11 ; Search time 104.01 Seconds
(without alignments)
9.980 Million cell updates/sec

Title: BASK-853-CLAIM4

Perfect score: 21

Sequence: 1 eagxxs 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

```

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	90.5	108	12	Q91TM3
2	19	90.5	122	17	Q9HPE3
3	19	90.5	130	15	Q36890
4	19	90.5	131	6	Q9GK47
5	19	90.5	143	17	Q9YBX6
6	19	90.5	145	16	Q9KRU8
7	19	90.5	149	10	Q9FRJ5
8	19	90.5	151	10	Q9AWI9
9	19	90.5	152	10	Q9XIK5
10	19	90.5	161	4	Q9BS09
11	19	90.5	161	4	Q9NWT1
12	19	90.5	164	2	Q54209
13	19	90.5	170	6	Q9MZ11
14	19	90.5	192	10	Q941T3
15	19	90.5	200	5	Q18144
16	19	90.5	201	5	O62323

17	19	90.5	204	16	O53374	053374 mycobacteri
18	19	90.5	212	16	Q92D00	Q92D00 listeria in
19	19	90.5	231	13	Q91306	Q91306 rana catesb
20	19	90.5	245	17	Q9HST1	Q9HST1 halobacteri
21	19	90.5	254	13	Q91307	Q91307 rana catesb
22	19	90.5	257	10	Q9XF64	Q9XF64 arabidopsis
23	19	90.5	257	10	Q9LZJ6	Q9LZJ6 arabidopsis
24	19	90.5	260	16	Q99RZ3	Q99RZ3 staphylococ
25	19	90.5	261	2	Q9ZIN7	Q9ZIN7 staphylococ
26	19	90.5	261	16	Q92NH7	Q92NH7 rhizobium m
27	19	90.5	266	12	Q88190	Q88190 soybean mos
28	19	90.5	267	12	Q88196	Q88196 soybean mos
29	19	90.5	269	11	Q9BCE4	Q9BCE4 mus musculu
30	19	90.5	270	5	Q9VSK4	Q9VSK4 drosophila
31	19	90.5	285	2	Q9AIT5	Q9AIT5 vibrio chol
32	19	90.5	287	13	Q93503	Q93503 xenopus lae
33	19	90.5	293	2	Q9L866	Q9L866 azospirillu
34	19	90.5	297	13	Q91296	Q91296 rana catesb
35	19	90.5	302	16	Q9PCJ6	Q9PCJ6 xylella fas
36	19	90.5	307	10	Q9FJG2	Q9FJG2 arabidopsis
37	19	90.5	311	13	Q90888	Q90888 gallus gall
38	19	90.5	311	13	Q90370	Q90370 coturnix co
39	19	90.5	313	13	Q9PUA6	Q9PUA6 xenopus lae
40	19	90.5	323	4	Q9Y5Q3	Q9Y5Q3 homo sapien
41	19	90.5	323	4	Q9H1F1	Q9H1F1 homo sapien
42	19	90.5	333	13	Q98TS3	Q98TS3 brachydanio
43	19	90.5	346	10	Q9ATS6	Q9ATS6 arundinella
44	19	90.5	350	16	Q981H9	Q981H9 rhizobium l
45	19	90.5	352	16	Q92WT2	Q92WT2 rhizobium m
46	19	90.5	356	13	Q98UK5	Q98UK5 brachydanio
47	19	90.5	356	13	Q73679	Q73679 brachydanio
48	19	90.5	357	10	O23101	O23101 arabidopsis
49	19	90.5	361	5	Q9BL91	Q9BL91 caenorhabdi
50	19	90.5	362	11	Q9JHQ1	Q9JHQ1 rattus norv

ALIGNMENTS

RESULT 1

```

Q91TM3 ID Q91TM3 PRELIMINARY; PRT; 108 AA.
AC Q91TM3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE T73.
OS Tupaia herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RA Bahr U., Darai G.;
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree
  Shrew) Herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Darai G., Bahr U.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF281817; AAK57118.1;
SQ SEQUENCE 108 AA; 11188 MW; 5732B2C61DBDE820 CRC64;

```

Query Match 90.5%; Score 19; DB 12; Length 108;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6

Db 30 EAGASS 35
||| I
RESULT 2
Q9HPE3 PRELIMINARY; PRT; 122 AA.
AC Q9HPE3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE VNG1678H.
DE VNG1678H.
GN Halobacterium sp. (strain NRC-1).
OS Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OC Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithacker B., Kreller K., Cruz R., Danson M.J., Hough D.W.,
RA MacDocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005076; AAG19927.1; -
KW Complete proteome.
SQ SEQUENCE 122 AA; 12001 MW; BE2432416C2BA256 CRC64;

Query Match 90.5%; Score 19; DB 17; Length 122;
Best Local Similarity 66.7%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 eagxxx 6
||| I
Db 57 EAGASS 62

RESULT 3
O36890 PRELIMINARY; PRT; 130 AA.
AC O36890;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG POLYPROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445059; PubMed=9300048;
RA Leigh Brown A.J., Lobidel D., Wade C.M., Rebus S., Phillips N.,
RA Brettell R.P., France A.J., Leen C.S., McMenamin J., McMillan A.,
RA Maw R.D., Mulcahy F., Robertson J.R., Sankar K.N., Scott G., Wyld R.,
RA Featherer J.F.;
RT "The molecular epidemiology of human immunodeficiency virus type 1 in
six cities in Britain and Ireland.";
RL Virology 235:166-177(1997).
DR EMBL; AF014287; AAC58368.1; -
DR InterPro; IPR000071; Retroviral_gag_pl7.
DR Pfam; PF00340; gag_pl7; 1.
DR PRINTS; PR00234; HIVIMATRIX.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
NON_TER 130 130

SQ SEQUENCE 130 AA; 14476 MW; 9053DD2EF3A3E00F CRC64;

Query Match 90.5%; Score 19; DB 15; Length 130;
Best Local Similarity 66.7%; Pred. No. 7.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 eagxxx 6
||| I
Db 99 EAGSS 104

RESULT 4
Q9GK47 PRELIMINARY; PRT; 131 AA.
AC Q9GK47;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RELAXIN-LIKE PROTEIN.
OS Galago crassicaudatus (Thick-tailed galago) (Otolemur crassicaudatus).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Otolemur.
OX NCBI_TaxID=9463;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Klonisch T., Froehlich C., Tetens F., Fischer B., Hombach-Klonisch S.;
RT "Molecular remodeling of members of the relaxin family during primate
evolution.";
RL Mol. Biol. Evol. 0:0-0(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; AF317624; AAG42317.1; -
DR InterPro; IPR000739; Insulin_IGF_relaxin.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 131 AA; 14414 MW; F18AFA9ACFC85943 CRC64;

Query Match 90.5%; Score 19; DB 6; Length 131;
Best Local Similarity 66.7%; Pred. No. 7.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 eagxxx 6
||| I
Db 54 EAGTSS 59

RESULT 5
Q9YBX6 PRELIMINARY; PRT; 143 AA.
AC Q9YBX6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 15.7 KDA PROTEIN APE1474.
GN APE1474.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kubota Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic

RT Crenarchaeon, Aeropyrum pernix K1.;"

RL DNA Res. 6:83-101(1999).

DR EMBL; AP000061; BAA80472.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 143 AA; 15708 MW; C40E29CBB0AF6892 CRC64;

Query Match 90.5%; Score 19; DB 17; Length 143;
Best Local Similarity 66.7%; Pred. No. 7.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 85 EAGAAS 90

RESULT 6

ID Q9KR08 PRELIMINARY; PRT; 145 AA.

AC Q9KR08

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL PROTEIN VC1536.

GN VC1536.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Elsen J.A., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P., McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";

RL Nature 406:477-483(2000).

RL EMBL; AE004231; AAF94690.1; -

DR TIGR; VC1536; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 145 AA; 15726 MW; 976EIF5EB50DB0EC CRC64;

Query Match 90.5%; Score 19; DB 16; Length 145;
Best Local Similarity 66.7%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 114 EAGSTS 119

RESULT 7

ID Q9FRJ5 PRELIMINARY; PRT; 149 AA.

AC Q9FRJ5

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 15.3 KDA PROTEIN.

GN OSJNB0064P21.9.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E., Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;

RT "Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence.";

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC073166; AAG46108.1; -

KW Hypothetical protein.

SQ SEQUENCE 149 AA; 15284 MW; 8B71E92310872766 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 149;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 111 EAGASS 116

RESULT 8

ID Q9AWI9 PRELIMINARY; PRT; 151 AA.

AC Q9AWI9

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE P0489A05.8 PROTEIN.

GN P0489A05.8.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0489A05.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP003105; BAB32988.1; -

SQ SEQUENCE 151 AA; 16103 MW; 6E942A203BC62C11 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 151;
Best Local Similarity 66.7%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |

Db 106 EAGAAS 111

RESULT 9

ID Q9XIK5 PRELIMINARY; PRT; 152 AA.

AC Q9XIK5

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE T10024.6.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Shinn P., Brooks S., Buehler E., Chao O., Dunn P., Khan S., Kim C., Walker M., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A., Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S.,

RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,
 RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC T10024 from Chromosome
 RT 1.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007067; AAD39566.1; -.
 SQ SEQUENCE 152 AA; 17676 MW; A7053F4DA73C3490 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 152;
 Best Local Similarity 66.7%; Pred. No. 8.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 |||||
 Db 138 EAGTTS 143

RESULT 10
 Q9BS09 ID Q9BS09 PRELIMINARY; PRT; 161 AA.
 AC Q9BS09;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 17.1 KDA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON ADENOCARCINOMA;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005805; AAH05805.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 161 AA; 17058 MW; E4098AB1F0A5D706 CRC64;

Query Match 90.5%; Score 19; DB 4; Length 161;
 Best Local Similarity 66.7%; Pred. No. 8.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 |||||
 Db 43 EAGSSS 48

RESULT 11
 Q9NW11 ID Q9NW11 PRELIMINARY; PRT; 161 AA.
 AC Q9NW11;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE CDNA FLJ20847 FIS, CLONE ADKA01746.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ADIPOSE TISSUE;
 RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
 RA Hiroo M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
 RA Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK000854; BAA91399.1; -.
 SQ SEQUENCE 161 AA; 17118 MW; 11098AB1EA15D71C CRC64;

Query Match 90.5%; Score 19; DB 4; Length 161;
 Best Local Similarity 66.7%; Pred. No. 8.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 |||||
 Db 43 EAGSSS 48

RESULT 12
 Q54209 ID Q54209 PRELIMINARY; PRT; 164 AA.
 AC Q54209;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE FABD, FABH, FABG, FABB, AND ORF5 GENES.
 OS Streptomyces glaucescens.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1907;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GLA.0;
 RX MEDLINE=95352622; PubMed=7626609;
 RA Summers R.G., Ali A., Shen B., Wessel W.A., Hutchinson C.R.;
 RT "Malonyl-coenzyme A:acyl carrier protein acyltransferase of
 RT Streptomyces glaucescens; a possible link between fatty acid and
 RT polyketide biosynthesis.";
 RL Biochemistry 34:9389-9402(1995).
 DR EMBL; LA3074; AAA99450.1; -.
 SQ SEQUENCE 164 AA; 18203 MW; CB0ECF031044BB09 CRC64;

Query Match 90.5%; Score 19; DB 2; Length 164;
 Best Local Similarity 66.7%; Pred. No. 9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 |||||
 Db 55 EAGTAS 60

RESULT 13
 Q9MZ11 ID Q9MZ11 PRELIMINARY; PRT; 170 AA.
 AC Q9MZ11;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 5-HT1A (FRAGMENT).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RA Pootanakit K., Hunter D.D., Brunken W.J.;
 RT "5-HT1A and 5-HT7 Receptor Expression in the Mammalian Retina.";
 RL Brain Res. 0:0-0(2000).
 DR EMBL; AF269231; AAF76184.1; -.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 FT NON_TER 1
 FT NON_TER 170 170
 SQ SEQUENCE 170 AA; 18518 MW; 42A5B4CF917B3250 CRC64;

Query Match 90.5%; Score 19; DB 6; Length 170;
 Best Local Similarity 66.7%; Pred. No. 9.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 142 EGAAS 147

RESULT 14

ID Q94LT3 PRELIMINARY; PRT; 192 AA.
AC Q94LT3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 19.8 KDA PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
RA Uterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNB0011A08 genomic sequence.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC034258; AAK54288.1; -;
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 19819 MW; 4CE8C88AE83DF374 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 192;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 40 EGAAS 45

RESULT 15

ID Q18144 PRELIMINARY; PRT; 200 AA.
AC Q18144;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 22.0 KDA PROTEIN.
GN C25A8.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA LaReille P., Stelliys L.;
RT "The sequence of C. elegans cosmid C25A8.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;

RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U61958; AAB03180.1; -;
KW Hypothetical protein.
SQ SEQUENCE 200 AA; 22012 MW; 66A23EDA709C66B2 CRC64;

Query Match 90.5%; Score 19; DB 5; Length 200;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 109 EGAAS 114

RESULT 16

ID O62323 PRELIMINARY; PRT; 201 AA.
AC O62323;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE R02D5.7 PROTEIN.
GN R02D5.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z78015; CAB01436.1; -;
SQ SEQUENCE 201 AA; 22266 MW; EC0423A8D7DDE4FE CRC64;

Query Match 90.5%; Score 19; DB 5; Length 201;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
||| |
Db 110 EGAAS 115

RESULT 17

ID O53374 PRELIMINARY; PRT; 204 AA.
AC O53374;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 21.6 KDA PROTEIN.
GN MOA3 OR RV3322C OR MTV016.22C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G., "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.", Nature 393:537-544(1998).	01-NOV-1996 (TrEMBLrel. 01, Created)
RL	Nature 393:537-544(1998).	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DR	EMBL: AL021841; CAA17094.1; --	MYOSIN VI (FRAGMENT).
DR	TubercuList; RV3322C; --	Rana catesbeiana (Bull frog).
DR	InterPro: IPR000051; SAM_bind.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
KW	Hypothetical protein; Complete proteome.	NCBI_TaxID=8400;
SQ	SEQUENCE 204 AA; 21614 MW; 13C5CB74C9C4B07F CRC64;	[1]
Query Match 90.5%; Score 19; DB 16; Length 204;		
Best Local Similarity 66.7%; Pred. No. 1.1e+03;		
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1 eagxxs 6	
Db	153 EAGTAS 158	
RESULT 18		
Q92D00		
ID	Q92D00 PRELIMINARY; PRT; 212 AA.	
AC	Q92D00;	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DE	LIN1021 PROTEIN.	
GN	LIN1021.	
OS	Listeria innocua.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;	
OC	Bacillus/Staphylococcus group; Listeria.	
OX	NCBI_TaxID=1642;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-CLIP 11262 / SEROVAR 5A;	
RX	PubMed-11679669;	
RA	Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A., Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T., Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P., Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O., Entian K.-D., Eshii H., Garcia-del Portillo F., Garfido P., Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D., Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G., Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H., Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R., Remmel B., Rose M., Schluter T., Simoes N., Tierrez A., Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.; "Comparative genomics of Listeria species.", Science 294:849-852(2001).	
RL	Science 294:849-852(2001).	
DR	EMBL: AL596167; CAC96252.1; --	
DR	ListiList; LIN1021; --	
KW	Complete proteome.	
SQ	SEQUENCE 212 AA; 23672 MW; CB73DB2965A08F99 CRC64;	
Query Match 90.5%; Score 19; DB 16; Length 212;		
Best Local Similarity 66.7%; Pred. No. 1.2e+03;		
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1 eagxxs 6	
Db	95 EAGASS 100	
RESULT 19		
Q91306		
ID	Q91306 PRELIMINARY; PRT; 231 AA.	
AC	Q91306;	
Query Match 90.5%; Score 19; DB 17; Length 245;		
Best Local Similarity 66.7%; Pred. No. 1.4e+03;		
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1 eagxxs 6	
Db	124 EAGSTS 129	

DT	01-NOV-1996 (TrEMBLrel. 01, Created)	
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE	MYOSIN VI (FRAGMENT).	
OS	Rana catesbeiana (Bull frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.	
OX	NCBI_TaxID=8400;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUP-SACCULE;	
RA	Solic C.F., Derflier B.H., Duyk G.M., Corey D.P.; "Molecular cloning of myosins from the bullfrog saccular macula: A candidate for the hair-cell adaptation motor.", Aud. Neurosci. 1:63-75(1994).	
RL	EMBL: U14380; AAA65089.1; --	
DR	HSP: P08799; INMD.	
DR	InterPro: IPR001609; myosin_head.	
DR	Pfam: PF00063; myosin_head; 2.	
DR	ProDom: PD000355; myosin_head; 1.	
FT	NON_TER 1	
FT	NON_TER 231	
SQ	SEQUENCE 231 AA; 25693 MW; D3FF5C343E6FAC8 CRC64;	
Query Match 90.5%; Score 19; DB 13; Length 231;		
Best Local Similarity 66.7%; Pred. No. 1.3e+03;		
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1 eagxxs 6	
Db	124 EAGSTS 129	
RESULT 20		
Q9HST1		
ID	Q9HST1 PRELIMINARY; PRT; 245 AA.	
AC	Q9HST1;	
DT	01-MAR-2001 (TrEMBLrel. 16, Created)	
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)	
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE	L-ISOASPARTYL PROTEIN CARBOXYL METHYLTRANSFERASE.	
GN	PLMT1 OR VNG0089G.	
OS	Halobacterium sp. (strain NRC-1).	
OC	Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;	
OC	Halobacterium.	
OX	NCBI_TaxID=64091;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE-20504483; PubMed-11016950;	
RA	Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J., Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W., Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.; "Genome sequence of Halobacterium species NRC-1.", Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).	
RL	EMBL: AE004977; AAG18721.1; --	
DR	InterPro: IPR000682; PCMT.	
DR	InterPro: IPR000051; SAM_bind.	
DR	Pfam: PF01135; PCMT; 1.	
KW	Transferase; Methyltransferase; Complete proteome.	
SQ	SEQUENCE 245 AA; 26216 MW; A26FBBBCFAA5DB78 CRC64;	
Query Match 90.5%; Score 19; DB 17; Length 245;		
Best Local Similarity 66.7%; Pred. No. 1.4e+03;		
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1 eagxxs 6	

Db 151 EAGAAS 156
||| |

RESULT 21

ID Q91307 PRELIMINARY; PRT; 254 AA.
AC Q91307;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MYOSIN VI (FRAGMENT).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCACULE;
RA SOLC C.F., Derfler B.H., Duyk G.M., Corey D.P.;
RT "Molecular cloning of myosins from the bullfrog sacculus macula: A candidate for the hair-cell adaptation motor.";
DR EMBL; U14381; AAA65090.1; -;
DR HSP; P10587; IBR2.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 3.
DR ProDom; PD000355; myosin_head; 1.
FT NON_TER 1
FT TER 254
SQ SEQUENCE 254 AA; 29039 MW; DB8395E86BD6DE93 CRC64;

Query Match 90.5%; Score 19; DB 13; Length 254;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
||| |

RESULT 22

ID Q9XF64 PRELIMINARY; PRT; 257 AA.
AC Q9XF64;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RING-H2 ZINC FINGER PROTEIN ATLS.
GN ATLS.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99408259; PubMed=10480382;
RX Salinas-Mondragon R.E., Garciduenas-Pina C., Guzman P.;
RT "Early elicitor induction in members of a novel multigene family coding for highly related RING-H2 proteins in Arabidopsis thaliana.";
FT Plant Mol. Biol. 40:579-590(1999).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AF132015; AAD33583.1; -;
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
KW Zinc-finger.
SQ SEQUENCE 257 AA; 28608 MW; 07BCEBF8CEC928C96 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 257;

Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
||| |

RESULT 23

ID Q9LZJ6 PRELIMINARY; PRT; 257 AA.
AC Q9LZJ6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RING-H2 ZINC FINGER PROTEIN ATLS.
GN F26K9_120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bloeker H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
RA Salanoubat M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AL162651; CAB83119.1; -;
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
KW Zinc-finger.
SQ SEQUENCE 257 AA; 28592 MW; B6B7595DFF528431 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 257;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxx 6
||| |

RESULT 24

ID Q99RZ3 PRELIMINARY; PRT; 260 AA.
AC Q99RZ3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MODA PROTEIN.
GN MODA OR SA2074.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Uji Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus

RT aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003136; BAB43371.1; -.
 RA HSSP; P37329; 1WOD.
 KW Complete proteome.
 SQ SEQUENCE 260 AA; 29117 MW; 7A5D4A01A4482C4D CRC64;

Query Match 90.5%; Score 19; DB 16; Length 260;
 Best Local Similarity 66.7%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 227 EAGATS 232

RESULT 25
 Q92IN7 PRELIMINARY; PRT; 261 AA.
 AC Q92IN7;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MODA.
 GN MODA.
 OS Staphylococcus carnosus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1281;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TM300;
 RX MEDLINE=98340502; PubMed=9675851;
 RA Neubauer H., Pantel I., Gotz F.;
 RT "Characterization of moeb-part of the molybdenum cofactor
 biosynthesis gene cluster in Staphylococcus carnosus.";
 RL FEMS Microbiol. Lett. 164:55-62(1998).
 DR EMBL; AF109295; AAC83133.1; -.
 DR HSSP; P37329; 1WOD.
 SQ SEQUENCE 261 AA; 29203 MW; 126A2D314BBAFB13 CRC64;

Query Match 90.5%; Score 19; DB 2; Length 261;
 Best Local Similarity 66.7%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 228 EAGATS 233

RESULT 26
 Q92NH7 PRELIMINARY; PRT; 261 AA.
 AC Q92NH7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.
 GN SMC01615.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21368234; PubMed=11474104;
 RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
 RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
 RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,

RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
 RA Ramsberger U., Surzycki R., Thebault P., Vandenbol M., Vandenbol M.,
 RA Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont sinorhizobium meliloti.";
 RL Science 293:668-672(2001).
 DR EMBL; AL591790; CAC46806.1; -.
 KW Complete proteome.
 SQ SEQUENCE 261 AA; 28115 MW; 031D03708E1084CB CRC64;

Query Match 90.5%; Score 19; DB 16; Length 261;
 Best Local Similarity 66.7%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 16 EAGTAS 21

RESULT 27
 Q88190 PRELIMINARY; PRT; 266 AA.
 ID Q88190;
 AC Q88190;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE COAT PROTEIN.
 OS Soybean mosaic virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=12222;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA;
 RA Chu R., Leng X., Bao Y., Pu Z., Pan N., Chen Z.;
 RT "Amplification of soybean mosaic virus coat protein gene by polymerase
 chain reaction and its sequence analysis.";
 RL Acta Bot. Sin. 34:523-528(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA;
 RA Xu L.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U25673; AAA70095.1; -.
 DR InterPro: IPR001592; Poty_coat.
 DR Pfam: PF00767; Poty_coat; 1.
 SQ SEQUENCE 266 AA; 30084 MW; 4E08AFE7D434307F CRC64;

Query Match 90.5%; Score 19; DB 12; Length 266;
 Best Local Similarity 66.7%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
 ||| |
 Db 25 EAGTSS 30

RESULT 28
 Q88196 PRELIMINARY; PRT; 267 AA.
 ID Q88196;
 AC Q88196;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 OS Soybean mosaic virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=12222;

RN SEQUENCE FROM N.A.
 RC STRAIN=CHINESE;
 RA Chu R.;
 RT "CDNA sequence of the gene encoding coat protein of SMV.";
 RL Nucleic Acids Res. 0:0-0(0).
 DR EMBL: X63771; CAA45307.1; -;
 DR InterPro: IPR001592; Poty_coat.
 FT Pfam: PF00767; Poty_coat; 1.
 FT NON_TER 1
 SQ SEQUENCE 267 AA; 30104 MW; 220E42F2595BE059 CRC64;

Query Match 90.5%; Score 19; DB 12; Length 267;
 Best Local Similarity 66.7%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6

Db 26 EAGTSS 31

RESULT 29

Q9DCE4 PRELIMINARY; PRT; 269 AA.
 AC Q9DCE4; 2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE O610039P13RIK PROTEIN.
 GN O610039P13RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK002854; BAB22409.1; -;
 DR MGD; MGI:1921346; O610039P13RIK.
 DR InterPro: IPR000636; Cation_chan_non_lig.
 SQ SEQUENCE 269 AA; 31242 MW; B549CB553DEB6568 CRC64;

Query Match 90.5%; Score 19; DB 11; Length 269;
 Best Local Similarity 66.7%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6

Db 71 EAGSAS 76

RESULT 30

Q9VGK4 PRELIMINARY; PRT; 270 AA.
 AC Q9VGK4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CG14714 PROTEIN.
 GN CG14714.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Pal,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Styrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003692; AAF54673.1; -;
 DR HSP; P28827; IRPM.
 DR FlyBase; FBgn0037929; CG14714.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR InterPro: IPR000242; Tyr_prot_phptase.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00194; PTPc; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
 KW Hydrolase.
 SQ SEQUENCE 270 AA; 30259 MW; 5C136F3135CAD001 CRC64;

Query Match

90.5%; Score 19; DB 5; Length 270;

Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 169 EAGSTS 174

RESULT 31
Q9AIT5 PRELIMINARY; PRT; 285 AA.
ID Q9AIT5;
AC Q9AIT5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ICMF-LIKE PROTEIN (FRAGMENT).
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=569B;
RA MEDLINE=20434574; PubMed=10981695;
RX Das S., Chakraborty A., Banerjee R., Roychoudhury S., Chaudhuri K.;
RT "Comparison of global transcription responses allows identification of
RT Vibrio cholerae genes differentially expressed following infection.";
RL FEMS Microbiol. Lett. 190:87-91(2000).
DR EMBL; AF339737; AAK27321.1; -.
FT NON_TER 1
FT NON_TER 285
SQ SEQUENCE 285 AA; 32165 MW; BDDE7FA9021F5661 CRC64;

Query Match 90.5%; Score 19; DB 2; Length 285;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 210 EAGSAS 215

RESULT 32
O93503 PRELIMINARY; PRT; 287 AA.
ID O93503;
AC O93503;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYRISTOYLATED ALANINE-RICH C KINASE SUBSTRATE.
GN MARCKS.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98030614; PubMed=9361009;
RA Shi Y., Sullivan S.K., Pitterle D.M., Kennington E.A., Graff J.M.,
RA Blackshear P.J.;
RT "Mechanisms of MARCKS gene activation during xenopus development.";
RL J. Biol. Chem. 272:29290-29300(1997).
DR EMBL; AF017299; AAC61897.1; -.
DR InterPro; IPR002101; MARCKS.
DR Pfam; PF02063; MARCKS; 1.
DR PRINTS; PR00963; MARCKS.
DR PROSITE; PS00826; MARCKS_1; 1.
DR PROSITE; PS00827; MARCKS_2; 1.
KW Kinase.
SQ SEQUENCE 287 AA; 29147 MW; 35CB7AE6090ED3C1 CRC64;

Query Match 90.5%; Score 19; DB 13; Length 287;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 110 EAGSTS 115

RESULT 33
Q9L866 PRELIMINARY; PRT; 293 AA.
ID Q9L866;
AC Q9L866;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 32.1 KDA PROTEIN.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OX Azospirillum.
RN NCBI_TaxID=192;
RX [1]
RP SEQUENCE FROM N.A.
RA Ma L., Zhao Y., Wang J., Li J.;
RT "Sequence and function analysis of draft genes downstream ORFs from
RT Azospirillum brasilense Yu62.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216815; AAF61911.1; -.
DR InterPro; IPR003310; DNA_glycosylase.
KW Hypothetical protein.
SQ SEQUENCE 293 AA; 32063 MW; 37417EA008F6BD61 CRC64;

Query Match 90.5%; Score 19; DB 2; Length 293;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 168 EAGAS 173

RESULT 34
Q91296 PRELIMINARY; PRT; 297 AA.
ID Q91296;
AC Q91296;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOSIN VI (FRAGMENT).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SACCULE;
RA Solc C.F., Derfler B.H., Duyk G.M., Corey D.P.;
RT "Molecular cloning of myosins from the bullfrog saccular macula: A
RT candidate for the hair-cell adaptation motor.";
RL Aud. Neurosci. 1:63-75(1994).
DR EMBL; U14370; AAA65079.1; -.
DR HSSP; P10587; IBR2.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 2.
DR ProDom; PD000355; myosin_head; 1.
FT NON_TER 1
FT NON_TER 297
SQ SEQUENCE 297 AA; 33755 MW; 46EE6C78A8ED530D CRC64;

Query Match 90.5%; Score 19; DB 13; Length 297;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;


```

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
   ||| |
Db 169 EAGSTS 174

RESULT 35
O9PCJ6 PRELIMINARY; PRT; 302 AA.
AC O9PCJ6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XF1783.
GN XF1783.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Hordin S., Bove J.M., Briones M.R.S.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Martins A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Maidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL; AE004000; AAF84591.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32047 MW; 758CC61DE4BE3590 CRC64;

Query Match 90.5%; Score 19; DB 16; Length 302;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
   ||| |
Db 106 EAGTAS 111

RESULT 36
O9FGJ2 PRELIMINARY; PRT; 307 AA.
AC O9FGJ2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

```

```

DE GB/AAD26962.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=COLUMBIA;
RX MEDLINE=99087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT Sequence features of the regions of 1,013,767 bp covered by sixteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:297-308(1998).
DR EMBL; AB015473; BAB08399.1; -
SQ SEQUENCE 307 AA; 35727 MW; 23CDB6C127CE90D1 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 307;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
   ||| |
Db 207 EAGTSS 212

RESULT 37
O90888 PRELIMINARY; PRT; 311 AA.
AC Q90888;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MAFB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95021288; PubMed=7935473;
RA Kataoka K., Fujiwara K.T., Noda M., Nishizawa M.;
RT "MafB, a new Maf family transcription activator that can associate with
RT Maf and Fos but not with Jun.";
RL Mol. Cell. Biol. 14:7581-7591(1994).
DR EMBL; D28600; BAA05938.1; -
DR InterPro; IPR001871; bZIP.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
SQ SEQUENCE 311 AA; 35467 MW; DDAE7F698B7D3ABA CRC64;

Query Match 90.5%; Score 19; DB 13; Length 311;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxs 6
   ||| |
Db 296 EAGSTS 301

RESULT 38
O90370 PRELIMINARY; PRT; 311 AA.
AC Q90370;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MAFB PROTEIN.

```

GN MAFB.
 OS *Coturnix coturnix japonica* (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96180718; PubMed=8620536;
 RA Sieweke M.H., Tekotte H., Frampton J., Graf T.;
 RT "WafB is an interaction partner and repressor of Ets-1 that inhibits
 erythroid differentiation.";
 RL Cell 85:49-60(1996).
 DR EMBL: X96511; CAA65360.1; -.
 DR InterPro: IPR001871; bZIP.
 DR Pfam: PF03131; bZIP_Maf; 1.
 DR SMART: SM00338; BRLZ; 1.
 SQ SEQUENCE 311 AA; 35476 MW; 7D1F3FA05D5CD683 CRC64;

Query Match 90.5%; Score 19; DB 13; Length 311;
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 296 EAGSTS 301

RESULT 39
 Q9PUA6 PRELIMINARY; PRT; 313 AA.
 AC Q9PUA6;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE bZIP TRANSCRIPTION FACTOR MAFB.
 GN MAFB.
 OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21152895; PubMed=11231068;
 RA Ishibashi S., Yasuda K.;
 RT "Distinct roles of maf genes during *Xenopus* lens development.";
 RL Mech. Dev. 101:155-166(2001).
 DR EMBL: AF202058; AAF08316.1; -.
 DR InterPro: IPR001871; bZIP.
 DR Pfam: PF03131; bZIP_Maf; 1.
 DR SMART: SM00338; BRLZ; 1.
 SQ SEQUENCE 313 AA; 35714 MW; 8E697A00A928BF95 CRC64;

Query Match 90.5%; Score 19; DB 13; Length 313;
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 298 EAGSTS 303

RESULT 40
 Q9V5Q3 PRELIMINARY; PRT; 323 AA.
 AC Q9V5Q3;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE MAFB/KREISLER BASIC REGION/LEUCINE ZIPPER TRANSCRIPTION FACTOR.

GN MAFB.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RX MEDLINE=99375320; PubMed=10444328;
 RA Wang P.W., Eisenbart J.D., Cordes S.P., Barsh G.S., Stoffel M.,
 RA Le Beau M.M.;
 RT "Human KRML (MAFB): cDNA cloning, genomic structure, and evaluation as
 a candidate tumor suppressor gene in myeloid leukemias.";
 RL Genomics 59:275-281(1999).
 DR EMBL: AF134157; AAD30106.1; -.
 DR InterPro: IPR001871; bZIP.
 DR Pfam: PF03131; bZIP_Maf; 1.
 DR SMART: SM00338; BRLZ; 1.
 SQ SEQUENCE 323 AA; 35829 MW; AE4DC23408E36E55 CRC64;

Query Match 90.5%; Score 19; DB 4; Length 323;
 Best Local Similarity 66.7%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 308 EAGSTS 313

RESULT 41
 Q9H1F1 PRELIMINARY; PRT; 323 AA.
 AC Q9H1F1;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE DJ644L1.1 (KREISLER (MOUSE) MAF-RELATED LEUCINE ZIPPER HOMOLOG).
 GN KRML.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ramsay H.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBDJ databases.
 DR EMBL: AL035665; CAB75863.1; -.
 DR InterPro: IPR001871; bZIP.
 DR SMART: SM00338; BRLZ; 1.
 SQ SEQUENCE 323 AA; 35792 MW; A0F3C09F8936CB16 CRC64;

Query Match 90.5%; Score 19; DB 4; Length 323;
 Best Local Similarity 66.7%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eaqxss 6
 |||||
 Db 308 EAGSTS 313

RESULT 42
 Q98TS3 PRELIMINARY; PRT; 333 AA.
 AC Q98TS3;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE HSP70 BINDING PROTEIN.
 OS *Brachydanio rerio* (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Ostariophysi;

OC Cypriniformes: Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Guerriero V., Raynes D.A.;
RT "Hsp70 binding protein from zebra fish (HspBPF).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY024336; AAG61257.1; -;
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 2.
DR SMART: SM00185; ARM; 2.
SQ SEQUENCE 333 AA; 37269 MW; E7C5ABD12F41D23E CRC64;

Query Match 90.5%; Score 19; DB 13; Length 333;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 23 EAGSAS 28

RESULT 43

Q9ATS6 PRELIMINARY; ERT; 346 AA.
AC Q9ATS6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TEOSONTE BRANCHED1 PROTEIN (FRAGMENT).
GN TBL.
OS Arundinella hirta.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Arundinelleae; Arundinella.
OX NCBI_TaxID=79825;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21165336; PubMed=11264415;
RA Lukens L., Doebley J.;
RT "Molecular evolution of the teosinte branched gene among maize and related grasses.";
RL Mol. Biol. Evol. 18:627-638(2001).
DR EMBL: AF322131; AAK37493.1; -;
FT NON_TER 346 1
FT NON_TER 346 346
SQ SEQUENCE 346 AA; 36902 MW; BFB5F29CD7449C89 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 346;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 280 EAGSAS 285

RESULT 44

Q981H9 PRELIMINARY; PRT; 350 AA.
AC Q981H9;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE MLR9369 PROTEIN.
GN MLR9369.
OS Rhizobium loti (Mesorhizobium loti).
OG Plasmid pMLa.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003016; BAB54976.1; -;
KW Plasmid; Complete proteome.
SQ SEQUENCE 350 AA; 39172 MW; BB075F9345BB9362 CRC64;

Query Match 90.5%; Score 19; DB 16; Length 350;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 108 EAGASS 113

RESULT 45

Q92WT2 PRELIMINARY; PRT; 352 AA.
AC Q92WT2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE ADENYLATE CYCLASE PROTEIN.
GN SMB20257.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL: AL603642; CAC48647.1; -;
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 352 AA; 37018 MW; F4AE6710196E06EF CRC64;

Query Match 90.5%; Score 19; DB 16; Length 352;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 13 EAGTSS 18

RESULT 46

Q98UK5 PRELIMINARY; PRT; 356 AA.
AC Q98UK5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TRANSCRIPTION FACTOR MAFB.
GN MAFB.

```

OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21064923; PubMed=11134968;
RA Kajihara M., Kawachi S., Kobayashi M., Ogino H., Takahashi S.,
RA Yasuda K.;
RT "Isolation, Characterization, and Expression Analysis of Zebrafish
RT Large Maf.";
RL J. Biochem. 129:139-146(2001).
DR EMBL: AB006322; BAB21102.1; -.
DR InterPro: IPR001871; bZIP.
DR Pfam: PF03131; bZIP_Maf; 1.
DR SMART: SM00338; BRLZ; 1.
SQ SEQUENCE 356 AA; 40233 MW; DE4C96B62C058865 CRC64;

Query Match 90.5%; Score 19; DB 13; Length 356;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
   ||| |
Db 341 EAGSTS 346

RESULT 47
O73679
ID 073679 PRELIMINARY; PRT; 356 AA.
AC 073679;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE TRANSCRIPTION FACTOR VAL.
GN VAL OR VALENTINO.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98165393; PubMed=9425134;
RA Moens C.B., Cordes S.P., Giorgianni M.W., Barsh G.S., Kimmel C.B.;
RT "Equivalence in the genetic control of hindbrain segmentation in fish
RT and mouse.";
RL Development 125:381-391(1998).
DR EMBL: AF006641; AAC18821.1; -.
DR ZFIN: ZDB-GENE-980526-515; val.
DR InterPro: IPR001871; bZIP.
DR Pfam: PF03131; bZIP_Maf; 1.
DR SMART: SM00338; BRLZ; 1.
SQ SEQUENCE 356 AA; 40243 MW; 07420DB0F6CD08F1 CRC64;

Query Match 90.5%; Score 19; DB 13; Length 356;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
   ||| |
Db 341 EAGSTS 346

RESULT 48
O23101
ID 023101 PRELIMINARY; PRT; 357 AA.
AC 023101;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE A_TM018A10.10 PROTEIN.
GN A_TM018A10.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA Dempsey S., Harper M.;
RT "The sequence of A. thaliana TM018A10.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA Washu;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF013294; AAB62869.1; -.
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR SMART: SM00256; FBOX; 1.
SQ SEQUENCE 357 AA; 40078 MW; B1683A07BF630633 CRC64;

Query Match 90.5%; Score 19; DB 10; Length 357;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
   ||| |
Db 91 EAGSSS 96

RESULT 49
Q9BL91
ID Q9BL91 PRELIMINARY; PRT; 361 AA.
AC Q9BL91;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 40.2 KDA PROTEIN.
GN Y18H1A.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Bemis G., Lamar B., Courtney L., Wohldmann P., Harrison M.;
RT "The sequence of C. elegans cosmid Y18H1A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";

```

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC024751; AAK21510.1; -
KW Hypothetical protein.
SQ SEQUENCE 361 AA; 40201 MW; 98F44C3F87D59625 CRC64;

Query Match 90.5%; Score 19; DB 5; Length 361;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 80 EAGSSS 85

RESULT 50
Q9JHQ1 PRELIMINARY; PRT; 362 AA.
AC Q9JHQ1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TITIN (FRAGMENT).
GN TTN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE=20490785; PubMed=11034912;
RA Person V., Kostin S., Suzuki K., Labelt S., Schaper J.;
RT "Antisense oligonucleotide experiments elucidate the essential role of
RT titin in sarcomerogenesis in adult rat cardiomyocytes in long-term
RT culture."
RL J. Cell Sci. 113:3851-3859(2000).
DR EMBL: AJ401157; CAB95001.1; -
DR HSP; P56276; ITLK.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; Igc2; 2.
DR SMART; SM00410; Ig_like; 1.
KW immunoglobulin domain.
FT NON_TER 1
FT NON_TER 362
SQ SEQUENCE 362 AA; 39601 MW; E8E6CE65BB7F4ED8 CRC64;

Query Match 90.5%; Score 19; DB 11; Length 362;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 eagxxx 6
||| |
Db 317 EAGSSS 322

Search completed: September 24, 2002, 11:26:14
Job time: 723 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:22:18 ; Search time 83.31 Seconds
(without alignments)
8.000 Million cell updates/sec

Title: BASK-853-CLAIM5
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : A_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	22	91.7	10	22 AAG94142	Human complementar
2	22	91.7	10	22 AAG97971	Human complementar
3	22	91.7	10	22 AAG97972	Human complementar
4	22	91.7	10	22 AAG97973	Human complementar
5	22	91.7	10	22 AAG97974	Human complementar
6	22	91.7	10	22 AAG97975	Human complementar
7	22	91.7	38	15 AAR58383	TSAR binding domai
8	22	91.7	50	22 AAU50435	Propionibacterium
9	22	91.7	52	21 AAB53513	Human colon cancer
10	22	91.7	53	22 AAU47468	Propionibacterium
11	22	91.7	58	22 AAU48691	Propionibacterium

12	22	91.7	59	22 AAO10898	Human polypeptide
13	22	91.7	61	22 AAU62078	Propionibacterium
14	22	91.7	62	22 AAU50474	Propionibacterium
15	22	91.7	76	22 AAU61979	Propionibacterium
16	22	91.7	78	20 AAY29072	T. gondii immunoge
17	22	91.7	78	22 AAU25543	T. gondii immunoge
18	22	91.7	80	22 AAU40314	Propionibacterium
19	22	91.7	82	22 AAU49223	Propionibacterium
20	22	91.7	86	22 AAG76165	Human colon cancer
21	22	91.7	90	22 AAM90260	Human immune/haema
22	22	91.7	91	22 ABB39057	Peptide #6563 enco
23	22	91.7	91	22 AAM95131	Human reproductive
24	22	91.7	91	22 AAM59709	Human brain expres
25	22	91.7	91	22 AAM72289	Human bone marrow
26	22	91.7	91	22 AAM32548	Peptide #6585 enco
27	22	91.7	106	22 AAM90809	Human immune/haema
28	22	91.7	109	21 AAG24569	Arabidopsis thalia
29	22	91.7	110	22 AAM92684	Human digestive sy
30	22	91.7	116	20 AAY12647	Human 5' EST secre
31	22	91.7	118	22 AAU65364	Propionibacterium
32	22	91.7	122	22 AAU50960	Propionibacterium
33	22	91.7	157	22 AAB92476	Human protein sequ
34	22	91.7	160	21 AAB32650	Eucalyptus grandis
35	22	91.7	160	21 AAG24568	Arabidopsis thalia
36	22	91.7	166	18 AAW32476	BBC6 protein for r
37	22	91.7	168	17 AAR95975	Nucleic acid recog
38	22	91.7	173	22 AAU38749	Novel subtilisin h
39	22	91.7	173	22 AAU38776	Novel subtilisin h
40	22	91.7	173	22 AAU38780	Novel subtilisin h
41	22	91.7	173	22 AAU38787	Novel subtilisin h
42	22	91.7	173	22 AAU38804	Novel subtilisin h
43	22	91.7	173	22 AAU38810	Novel subtilisin h
44	22	91.7	173	22 AAU38835	Novel subtilisin h
45	22	91.7	173	22 AAU38843	Novel subtilisin h
46	22	91.7	173	22 AAU38869	Novel subtilisin h
47	22	91.7	175	22 ABB68003	Drosophila melanog
48	22	91.7	182	22 ABB68002	Drosophila melanog
49	22	91.7	190	22 AAM40840	Human polypeptide
50	22	91.7	194	22 AAU40065	Propionibacterium

ALIGNMENTS

RESULT 1
AAG94142
ID AAG94142 standard; Peptide; 10 AA.
XX
AC AAG94142;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 336.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-G804776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides

PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
XX drug candidates or pro-drugs -

PS Example 4; Page 88; 646pp; English.

XX The invention relates to a set of complementary peptide ligands
XX generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match 91.7%; Score 22; DB 22; Length 10;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 5 gtssps 10

RESULT 2

AAG97971
ID AAG97971 standard; Peptide; 10 AA.

XX AC AAG97971;

XX DT 18-SEP-2001 (first entry)

XX DE Human complementary peptide, SEQ ID NO: 4166.

XX KW Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

XX PN WO200142277-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB04776.

XX PR 13-DEC-1999; 99GB-0029464.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
XX drug candidates or pro-drugs -

PS Example 6; Page 630; 646pp; English.

XX The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match 91.7%; Score 22; DB 22; Length 10;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 5 gtssps 10

RESULT 3

AAG97972
ID AAG97972 standard; Peptide; 10 AA.

XX AC AAG97972;

XX DT 18-SEP-2001 (first entry)

XX DE Human complementary peptide, SEQ ID NO: 4167.

XX KW Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

XX PN WO200142277-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB04776.

XX PR 13-DEC-1999; 99GB-0029464.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
XX drug candidates or pro-drugs -

PS Example 6; Page 630; 646pp; English.

XX The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match 91.7%; Score 22; DB 22; Length 10;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 5 gtssps 10

RESULT 4

AAG97973
ID AAG97973 standard; Peptide; 10 AA.

XX AC AAG97973;

XX DT 18-SEP-2001 (first entry)

XX DE Human complementary peptide, SEQ ID NO: 4168.

XX KW Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

Db 4 gtssps 9

RESULT 7

ID AAR58383 standard; Protein; 38 AA.

AC AAR58383;

DT 12-APR-1995 (first entry)

XX TSAR binding domain encoded by clone ZnlA7, binds zinc.

XX TSAR: totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; zinc binding.

XX Synthetic.

XX WO9418318-A.

XX 18-AUG-1994.

XX 01-FEB-1994; 94WO-US00977.

XX 01-FEB-1993; 93US-0013416.

XX 30-DEC-1993; 93US-0176500.

XX 31-JAN-1994; 94US-0189331.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Fowlkes DM, Kay BK;

XX WPI; 1994-279739/34.

XX N-PSDB; AAQ70472.

XX Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain

XX Claim 38; Page 95; 255pp; English.

XX AAR58376-94 show the amino acid sequences of the binding domain of TSAR
 CC (Totally Synthetic Affinity Reagents) peptides from the TSAR-9 library.
 CC These particular examples bind zinc. The non-variable amino acids at the
 CC NH2 and COOH terminals are not shown. TSAR peptides are generated using
 CC generic oligonucleotides (see AAQ70470-73 for examples). TSARs are
 CC concatenated heterofunctional proteins or peptides, comprising at least
 CC two functional regions - a binding domain with affinity for a ligand and
 CC a second effector peptide portion that is chemically or biologically
 CC active. They may further comprise a linker peptide between the 2 domains.
 CC The TSARs or comps. comprising a TSAR binding domain can be used in
 CC vivo to deliver a chemically or biologically active moiety, eg. metal
 CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
 CC on the cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed
 CC activity allowing direct and rapid detection in a screening process.

XX Sequence 38 AA;

Query Match 91.7%; Score 22; DB 15; Length 38;

Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

Db 20 gtssps 25

RESULT 8

AAU50435

ID AAU50435 standard; Protein; 50 AA.

XX AAU50435;

DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #11331.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59548.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX Example 1; SEQ ID No 11630; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).

XX Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 50 AA;

Query Match 91.7%; Score 22; DB 22; Length 50;

Best Local Similarity 66.7%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

Db 15 gtatps 20

XX	AAU47468;
AC	27-FEB-2002 (first entry)
DT	Propionibacterium acnes immunogenic protein #8364.
XX	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
DE	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX	dermatological; osteopathic; neuroprotectant.
XX	Propionibacterium acnes.
OS	WO200181581-A2.
PN	01-NOV-2001.
XX	20-APR-2001; 2001WO-US12865.
Pf	21-APR-2000; 2000US-199047P.
XX	02-JUN-2000; 2000US-208841P.
PR	07-JUL-2000; 2000US-216747P.
XX	(CORI-) CORIXA CORP.
PA	Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX	WPI; 2001-616774/71.
DR	N-PSDB; AAS59539.
XX	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris -
XX	Example 1; SEQ ID No 8663; 1069pp; English.
PS	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	downregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA).
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	Sequence 53 AA;
SQ	
Query Match	91.7%; Score 22; DB 22; Length 53;
Best Local Similarity	66.7%; Pred. No. 2.1e+02;
Matches	4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 gtxxps 6
Db	12 gtatsps 17
RESULT 11	
AAU48691	

ID XX AAU48691 standard; Protein; 58 AA.
 AC AAU48691;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #9587.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 9886; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 58 AA;

 Query Match 91.7%; Score 22; DB 22; Length 58;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 gtxxps 6
 Db 6 gtaxps 11

 RESULT 12
 ID AAU48691 standard; Protein; 58 AA.
 AC AAU48691;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #9587.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 9886; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 58 AA;

 Query Match 91.7%; Score 22; DB 22; Length 58;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 gtxxps 6
 Db 6 gtaxps 11

 RESULT 12
 ID AAU48691 standard; Protein; 58 AA.
 AC AAU48691;
 XX
 DT 27-FEB-2002 (first entry)
 XX

AA010898
 ID AA010898 standard; Protein; 59 AA.
 XX
 AC AA010898;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 24790.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-514838/56.
 DR N-PSDB; AA190829.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20; SEQ ID NO 24790; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and/or
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 59 AA;

 Query Match 91.7%; Score 22; DB 22; Length 59;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 gtxxps 6
 Db 49 gtttss 54

 RESULT 13
 ID AAU62078 standard; Protein; 61 AA.
 AC AAU62078;
 XX
 DT 27-FEB-2002 (first entry)
 XX

```

DE XX Propionibacterium acnes immunogenic protein #22974.
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
OS Propionibacterium acnes.
PN WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX
XX 02-JUN-2000; 2000US-208841P.
XX
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX
XX N-PSDB; AAS59623.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 23273; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 61 AA;

Query Match 91.7%; Score 22; DB 22; Length 61;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
Db 42 gtsps 47

RESULT 14
AAU50474
ID AAU50474 standard; Protein: 62 AA.
XX
XX AAU50474;
XX
XX 27-FEB-2002 (first entry)

DE XX Propionibacterium acnes immunogenic protein #11370.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX
XX 02-JUN-2000; 2000US-208841P.
XX
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX
XX N-PSDB; AAS59548.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 11669; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 62 AA;

Query Match 91.7%; Score 22; DB 22; Length 62;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
Db 43 gtstps 48

RESULT 15
AAU61979
ID AAU61979 standard; Protein: 76 AA.
XX
XX AAU61979;
XX

```

```

XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #22875.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX FN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX PR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59621.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX PS Example 1; SEQ ID No 23174; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 76 AA;

Query Match 91.7%; Score 22; DB 22; Length 76;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 2 gttssps 7

RESULT 16
AAAY29072
ID AAAY29072 standard; Protein; 78 AA.
XX
XX AC AAAY29072;

Query Match 91.7%; Score 22; DB 22; Length 76;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 2 gttssps 7

RESULT 17
AAU25543
ID AAU25543 standard; Protein; 78 AA.
XX
XX AC AAU25543;
XX DT 17-DEC-2001 (first entry)
XX DE T. gondii immunogenic protein PM2A18-a.
XX KW Immunogenic protein; oocyst; faeces; enteric apicomplexa oocyst;
XX KW Cryptosporidium oocyst; Toxoplasma oocyst; Giardia cyst; vaccine;
XX KW oocyte shedding.
XX OS Toxoplasma gondii.
XX FN US2001014447-A1.
XX PD 16-AUG-2001.
XX PR
XX AC

```

PF 18-DEC-1998; 98US-0216393.
 XX
 PR 19-DEC-1997; 97US-0994825.
 XX
 PA (MILH/) MILHAUSEN M J.
 XX
 PI Milhausen MJ;
 XX
 DR WPI; 2001-529100/58.
 DR N-PSDB; AAS42721.

XX Detecting parasite oocysts or cysts in faeces, comprises eluting DNA
 CC from sample into aqueous solution by heating, amplifying DNA with
 CC primers specific for oocysts of cysts being detected, and detecting
 CC amplification product -

PS Disclosure; Page 150; 188pp; English.

XX The invention relates to detection of parasite oocysts or cysts in
 CC a faeces sample comprising contacting the sample with a solid support,
 CC drying and then washing the sample with an aqueous wash solution, adding
 CC an aqueous elution solution and eluting DNA from the sample by heating
 CC and amplifying by PCR oocyst/cyst-specific DNA and detecting the
 CC amplification products. The method is useful for detecting parasite
 CC oocysts e.g., enteric apicomplexa oocysts such as Cryptosporidium oocysts
 CC or Toxoplasma oocysts, or for detecting parasite cysts e.g. Giardia
 CC cysts. The method is also useful for developing vaccines to prevent
 CC oocyte shedding in cats. The present sequence represents an
 CC immunogenic protein from Toxoplasma gondii.

XX Sequence 78 AA;

Query Match 91.7%; Score 22; DB 22; Length 78;
 Best Local Similarity 66.7%; Pred. NO. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 50 gtsps 55

RESULT 18
 AAU40914
 ID AAU40914 standard; Protein; 80 AA.
 XX
 AC AAU40914;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #1810.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59513.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX Example 1; SEQ ID No 2109; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, linitis,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 80 AA;

Query Match 91.7%; Score 22; DB 22; Length 80;
 Best Local Similarity 66.7%; Pred. NO. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 22 gtsps 27

RESULT 19
 AAU49223
 ID AAU49223 standard; Protein; 82 AA.

XX AAU49223;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #10119.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59545.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 10418; 1069pp; English.
PS
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, the central
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 82 AA;

Query Match 91.7%; Score 22; DB 22; Length 82;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
|| ||
Db 33 gttaps 38

RESULT 20
AAG76165
ID AAG76165 standard; Protein; 86 AA.
XX
AC AAG76165;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:6929.
DE
DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 12.
KW
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI WPI; 2001-235357/24.
XX
DR

DR N-PSDB; AAH35570.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 11; Page 8369; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAH77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 86 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
|| ||
Db 37 gttaps 42

RESULT 21
AAM90260
ID AAM90260 standard; Protein; 90 AA.
XX
AC AAM90260;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:17853.
DE
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
KW
OS Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 03-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233663.
PR 14-SEP-2000; 2000US-0233664.
PR 14-SEP-2000; 2000US-0233665.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

N-PSDB; AAK63041.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Claim 11; SEQ ID NO 17853; 307lpp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I)

CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to express the
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 90 AA;

Query Match 91.7%; Score 22; DB 22; Length 90;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 1 gtasps 6

RESULT 22

ABB39057
ID ABB39057 standard; Peptide; 91 AA.

XX AC ABB39057;

XX XX 04-FEB-2002 (first entry)

XX DE Peptide #6563 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX XX 09-AUG-2001.

XX XX 30-JAN-2001; 2001WO-US00669.

XX XX 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human fetal liver -

XX PS Claim 27; SEQ ID NO 31692; 639pp + sequence listing; English.

XX CC The invention relates to a single exon nucleic acid probe for

XX CC measuring human gene expression in a sample derived from human foetal

XX CC liver. The single exon nucleic acid probes may be used for predicting,

XX CC measuring and displaying gene expression in samples derived from human

XX CC foetal liver. The present sequence is a peptide encoded by a single exon

XX CC nucleic acid probe of the invention.

XX CC Note: The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX XX Sequence 91 AA;

Query Match 91.7%; Score 22; DB 22; Length 91;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 56 gtssps 61

RESULT 23

AAM95131

ID AAM95131 standard; Protein; 91 AA.

XX AC AAM95131;

XX XX 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen SEQ ID NO: 3789.

XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KW cancer; gene therapy.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX XX 02-AUG-2001.

XX XX 17-JAN-2001; 2001WO-US01339.

XX XX 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 11-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 14-JUL-2000; 2000US-0217496.

XX PR 26-JUL-2000; 2000US-0218290.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 14-AUG-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225447.

XX PR 14-AUG-2000; 2000US-0225757.

XX PR 14-AUG-2000; 2000US-0225758.

XX PR 18-AUG-2000; 2000US-0226279.

XX PR 22-AUG-2000; 2000US-0226681.

XX PR 22-AUG-2000; 2000US-0226868.

XX PR 23-AUG-2000; 2000US-0227182.

XX PR 23-AUG-2000; 2000US-0227009.

XX PR 30-AUG-2000; 2000US-0228924.

XX PR 01-SEP-2000; 2000US-0229287.

XX PR 01-SEP-2000; 2000US-0229343.

XX PR 01-SEP-2000; 2000US-0229344.

XX PR 05-SEP-2000; 2000US-0229509.

XX PR 05-SEP-2000; 2000US-0229513.

XX PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231415.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
DR N-PSDB; AAL01101.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
XX Claim 11; SEQ ID NO 3789; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a protein of the invention.
XX
SQ Sequence 91 AA;

Query Match 91.7%; Score 22; DB 22; Length 91;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 16 gtstps 21

RESULT 24
AAM59709
ID AAM59709 standard; Protein; 91 AA.
XX
XX AAM59709;
XX
XX 05-NOV-2001 (first entry)
DT
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 31814.
DE
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00667.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR

PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains .
 XX
 XX Example 4; SEQ ID NO: 31814; 650pp + Sequence Listing; English.
 PS The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 XX Sequence 91 AA;
 SQ
 Query Match 91.7%; Score 22; DB 22; Length 91;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 || ||
 Db 56 gtsps 61
 RESULT 25
 AAM72289
 ID AAM72289 standard; Protein; 91 AA.
 XX
 XX AAM72289;
 XX
 XX 06-NOV-2001 (first entry)
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 32595.
 DE
 DE Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 KW
 XX Homo sapiens.
 OS
 XX WO200157276-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00668.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow .
 XX
 XX Example 4; SEQ ID NO: 32595; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 XX Sequence 91 AA;
 SQ
 Query Match 91.7%; Score 22; DB 22; Length 91;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 || ||
 Db 56 gtsps 61
 RESULT 26
 AAM32548
 ID AAM32548 standard; Protein; 91 AA.
 XX
 XX AAM32548;
 XX
 XX 17-OCT-2001 (first entry)
 XX
 DE Peptide #6585 encoded by probe for measuring placental gene expression.
 XX
 XX Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 KW
 XX Homo sapiens.
 OS
 XX WO200157272-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00663.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488997/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta .
 PT
 XX
 PS Claim 27; SEQ ID No 32817; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 XX Sequence 91 AA;
 SQ

Query Match 91.7%; Score 22; DB 22; Length 91;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 56 gtsps 61

RESULT 27
AAM90809
ID AAM90809 standard; Protein; 105 AA.
XX
AC AAM90809;
XX
DT 07-NOV-2001 (first entry)
XX
XX Human Immune/haematopoietic antigen SEQ ID NO:18402.
DE
XX
KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
DR N-PSDB; AAK63590.
DR
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Claim 11; SEQ ID NO 18402; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 106 AA;

Query Match 91.7%; Score 22; DB 22; Length 106;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 37 gtsaps 42

RESULT 28
AAG24569
ID AAG24569 standard; Protein; 109 AA.
XX
AC AAG24569;
XX
XX 17-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28293.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142377.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 91.7%; Score 22; DB 21; Length 109;
Best Local Similarity 66.7%; Pred. NO. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
| | | |
Db 71 gtttpps 76

RESULT 29

AAM92684

ID AAM92684 standard; Protein; 110 AA.

XX AAM92684;

XX 06-NOV-2001 (first entry)

XX Human digestive system antigen SEQ ID NO: 2033.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX Homo sapiens.

```
XX WO200155314-A2.
PN 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01324.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX XX
```


PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-502630/55.
 DR N-PSDB; AAK88457.
 XX
 PT Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases -
 XX
 PS Claim 11; SEQ ID NO 2033; 986pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a digestive system antigen of
 CC the invention.
 XX
 SQ Sequence 110 AA;
 CC
 Query Match 91.7%; Score 22; DB 22; Length 110;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gtxxps 6
 DB 13 gtsaps 18
 || ||
 || ||
 RESULT 30
 AAY12647
 ID AAY12647 standard; Protein; 116 AA.
 XX
 AC AAY12647;
 XX
 XX 22-JUN-1999 (first entry)
 DT
 XX
 DE Human 5' EST secreted protein SEQ ID NO: 312 from WO 9906553.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.
 XX
 OS Homo sapiens.
 XX
 PN WO9906553-A2.
 XX
 PD 11-FEB-1999.
 XX
 XX 31-JUL-1998; 98WO-IB01237.
 XX
 XX 01-AUG-1997; 97US-0905051.
 XX
 PA (GEST) GENSET.
 XX
 XX Duclert A, Dumas Milne Edwards J, Lacroix B;
 PI
 XX WPI; 1999-153783/13.
 DR
 DR N-PSDB; AAX41505.
 XX
 XX New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from umbilical cord, lymph ganglia,
 PT lymphocytes and placental tissue
 XX
 PS Claim 34; Page 397-398; 411pp; English.
 XX
 CC AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12521 to

CC AAY12668, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, antiinflammatory
 CC thrombolytic activity, receptor/ligand activity, antiinflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 SQ Sequence 116 AA;
 CC
 Query Match 91.7%; Score 22; DB 20; Length 116;
 Best Local Similarity 66.7%; Pred. No. 4.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gtxxps 6
 DB 53 gtsaps 58
 || ||
 || ||
 RESULT 31
 AAU65364
 ID AAU65364 standard; Protein; 118 AA.
 XX
 AC AAU65364;
 XX
 XX 27-FEB-2002 (first entry)
 DT
 XX
 DE Propionibacterium acnes immunogenic protein #26260.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 XX 01-NOV-2001.
 PD
 XX
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX 21-APR-2000; 2000US-199047P.
 PR
 XX 02-JUN-2000; 2000US-208841P.
 PR
 XX 07-JUL-2000; 2000US-216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 PI
 XX WPI; 2001-616774/71.
 DR
 DR N-PSDB; AAS59666.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 26559; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 118 AA;

Query Match 91.7%; Score 22; DB 22; Length 118;
 Best Local Similarity 66.7%; Pred. No. 4.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gtxxps 6
 Db 111 gatatp 116

RESULT 32
 AAU50960
 ID AAU50960 standard; Protein; 122 AA.

XX AC AAU50960;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #11856.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI: 2001-616774/71.

XX DR N-PSDB; AAS59549.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for

XX PT vaccinating against and diagnosing infections, especially useful for

XX PT treating acne vulgaris -

XX PS Example 1; SEQ ID No 12155; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX CC polypeptides. The proteins and their associated DNA sequences are used in

XX CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 122 AA;

Query Match 91.7%; Score 22; DB 22; Length 122;
 Best Local Similarity 66.7%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gtxxps 6
 Db 18 gtttts 23

RESULT 33
 AAB92476
 ID AAB92476 standard; Protein; 157 AA.

XX AC AAB92476;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:10552.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI: 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs -

XX PS Claim 8; SEQ ID 10552; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602

XX CC full-length cDNAs defined in the specification. Where a primer set

XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC in gene therapy. The primer sets can be used in antisense therapy and
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX SQ Sequence 157 AA;

Query Match 91.7%; Score 22; DB 22; Length 157;
 Best Local Similarity 66.7%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 76 gtaaps 81

RESULT 34

AAH32650
 ID AAB32650 standard; Protein; 160 AA.

XX AC AAB32650;

XX DT 25-JAN-2001 (first entry)

XX DE Eucalyptus grandis transcription factor protein sequence #108.

XX KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB.

XX OS Eucalyptus grandis.

XX PN W0200053724-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06112.

XX PR 11-MAR-1999; 99US-0266513.

XX PR 18-AUG-1999; 99US-0149485.

XX XX (GENE-) GENESIS RES & DEV CORP LTD.
 (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX PI Wood M, McGrath A, Shenk MA, Glenn M;

XX DR WPI; 2000-579369/54.

XX XX New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide

XX PS Claim 8; Page 243; 747pp; English.

XX XX

CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
 CC zipper, LIM domain, AP2 and EREBS, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB.

XX SQ Sequence 160 AA;

Query Match 91.7%; Score 22; DB 21; Length 160;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 112 gtaaps 117

RESULT 35

AAG24568
 ID AAG24568 standard; Protein; 160 AA.

XX AC AAG24568;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28292.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135323.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 91.7%; Score 22; DB 21; Length 160;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 122 gtttpps 127

RESULT 36
 AAW32476
 ID AAW32476 standard; Protein; 165 AA.

AC AAW32476;

XX 15-JAN-1998 (first entry)

XX BBC6 protein for regulating cell death.

XX BBC6 gene; cell death; cell cycle; Bcl2; human.

XX Homo sapiens.

XX US5663316-A.

XX 02-SEP-1997.

XX 18-JUN-1996; 96US-0665617.

XX 18-JUN-1996; 96US-0665617.

XX (CLON-) CLONTECH LAB INC.

PI Xudong Y;

DR WPI: 1997-447980/41.

DR N-PSDB; AAT91561.

XX Isolated BBC6 gene - encodes a protein that regulates cell death through interaction with Bcl-2

XX Claim 1; Column 11-12; 7pp; English.

XX The present sequence represents a protein of 166 amino acids. The sequence is disclosed as being a protein called BBC6 which regulates cell death through interaction with Bcl-2. The DNA may be used for the production of the recombinant protein, which can be used in unspecified therapeutic or diagnostic procedures, as a molecular weight marker, and to raise antibodies that can be used in unspecified diagnostic or therapeutic applications and to reduce or eliminate the biological activity of the BBC6 protein in vivo.

XX Sequence 166 AA;

Query Match 91.7%; Score 22; DB 18; Length 166;
 Best Local Similarity 66.7%; Pred. No. 6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 160 gtaaps 165

RESULT 37
 AAU38749

AAR95975
 ID AAR95975 standard; peptide; 168 AA.

XX AAR95975;

XX 19-FEB-1997 (first entry)

XX Nucleic acid recognition unit #11.

XX Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SP1; TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus; HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence; nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria; virus.

XX Synthetic.

XX WO9617956-A2.

XX 13-JUN-1996.

XX 07-DEC-1995; 95WO-US15944.

XX 09-DEC-1994; 94US-0353476.

XX (GENE-) GENE POOL INC.

XX Weininger AM, Weininger S;

XX WPI: 1996-287199/29.

XX Probe nucleic acids, target binding assemblies, etc - for detection and localisation of specific nucleic acid sequences, esp. HIV and HPV

XX Claim 14; Page 97-98; 172pp; English.

XX AAR95965-R95993 represent the nucleic acid recognition units (NAR) of target binding assemblies (TBA) of the invention. These NARs are selected from NF-kappa-B, SP1, TATA, human papillomavirus (HPV) E2, HPV LTR, human immunodeficiency virus (HIV) LTR and Tat binding units. This sequence represents a SP1 sequence NAR. The TBA is recognised by the target binding region (TBR) of a probe of the invention. The probe of the invention contains a TBR, a booster binding region (BBR), and an optional support or attachment (OSA). The TBA contains at least one NAR, and optionally a linker sequence, an assembly sequence, an asymmetry sequence, a nuclear localisation signal sequence, and an OSA. The assembly sequence and asymmetry sequences are responsible for the folding and association of the NARs. The linker sequence is an oligopeptide, which does not interfere with NAR function, but provides stability and control over the spacing of the NAR from the rest of the TBA. The OSA is an attached support or indicator, or other means of localisation of the probe. The probe can be used in a method for detecting or localising a specific target nucleic acid sequence (TNA). The method is highly sensitive, and has a high degree of specificity. The method can be used for detecting specific nucleic acid sequences, including those found in human cells, in HIV, HPV, and other nucleic acid containing systems, including bacteria and viruses.

XX Sequence 168 AA;

Query Match 91.7%; Score 22; DB 17; Length 168;
 Best Local Similarity 66.7%; Pred. No. 6.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 119 gttatps 124

RESULT 38
 AAU38749

ID XX AAU38749 standard; Protein; 173 AA.
AC AAU38749;
DT 19-DEC-2001 (first entry)
DE Novel subtilisin homologue #6.
KW Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
KW contact lens cleansing solution; dry cleaning.
OS Bacillus sp.
FN WO200175087-A2.
XX 11-OCT-2001.
XX 02-APR-2001; 2001WO-US10781.
XX 03-APR-2000; 2000US-194143P.
XX (MAXY-) MAXYGEN INC.
XX Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
PI Minshall J;
XX WPI; 2001-616689/71.
DR N-PSDB; AAS56604.
XX Novel subtilisin homologue polypeptides having improved endo-protease
PT activity relative to mature Savinase subtilisin polypeptide, useful as
PT component of cleaning solutions e.g. laundry detergents, dry cleaning
XX
PS Claim 1; Page 103; 138pp; English.
XX The invention relates to an isolated polypeptide with improved endo-
CC protease activity relative to subtilisin homologue polypeptide, Savinase
CC (RTM). The invention also relates to an integrated system comprising a
CC computer or computer readable medium comprising a database comprising one
CC or more sequence records. Each record comprises one or more character
CC string corresponding to a nucleic acid or protein sequence of AAS56599-
CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
CC comprises a user input interface allowing a user to selectively
CC one or more sequence record. The integrated system is useful for
CC presenting information pertaining to one of several sequence records
CC stored in a database. The method involves determining a list of one or
CC more character strings corresponding to the sequence or its subsequence,
CC determining which character strings of the list are selected by a user,
CC and displaying the selected character strings or aligning the selected
CC character string with an additional character string. The method
CC further involves displaying an alignment of the selected character
CC string with the additional character string, and displaying the list.
CC The polynucleotides are useful as probes, primers, sense and antisense
CC nucleotides. The subtilisin homologue polypeptides are useful for
CC producing antibodies which have diagnostic uses related to the activity,
CC distribution and expression of subtilisin homologues. The subtilisin
CC homologue polypeptides are used in compositions that serve as cleaning
CC solutions in a wide variety of applications including laundry detergents,
CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
CC represent the amino acid sequences of novel subtilisin homologues of the
CC invention.
XX
SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 71 gtsaps 76

RESULT 39
AAU38776
ID AAU38776 standard; Protein; 173 AA.
XX
AC AAU38776;
XX
DT 19-DEC-2001 (first entry)
XX
DE Novel subtilisin homologue #33.
XX
KW Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
KW contact lens cleansing solution; dry cleaning.
OS Bacillus sp.
XX
XX WO200175087-A2.
XX 11-OCT-2001.
XX 02-APR-2001; 2001WO-US10781.
XX 03-APR-2000; 2000US-194143P.
XX (MAXY-) MAXYGEN INC.
XX Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
PI Minshall J;
XX WPI; 2001-616689/71.
DR N-PSDB; AAS56631.
XX
XX Novel subtilisin homologue polypeptides having improved endo-protease
PT activity relative to mature Savinase subtilisin polypeptide, useful as
PT component of cleaning solutions e.g. laundry detergents, dry cleaning
XX
PS Claim 1; Page 106; 138pp; English.
XX The invention relates to an isolated polypeptide with improved endo-
CC protease activity relative to subtilisin homologue polypeptide, Savinase
CC (RTM). The invention also relates to an integrated system comprising a
CC computer or computer readable medium comprising a database comprising one
CC or more sequence records. Each record comprises one or more character
CC string corresponding to a nucleic acid or protein sequence of AAS56599-
CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
CC comprises a user input interface allowing a user to selectively
CC one or more sequence record. The integrated system is useful for
CC presenting information pertaining to one of several sequence records
CC stored in a database. The method involves determining a list of one or
CC more character strings corresponding to the sequence or its subsequence,
CC determining which character strings of the list are selected by a user,
CC and displaying the selected character strings or aligning the selected
CC character string with an additional character string. The method
CC further involves displaying an alignment of the selected character
CC string with the additional character string, and displaying the list.
CC The polynucleotides are useful as probes, primers, sense and antisense
CC nucleotides. The subtilisin homologue polypeptides are useful for
CC producing antibodies which have diagnostic uses related to the activity,
CC distribution and expression of subtilisin homologues. The subtilisin
CC homologue polypeptides are used in compositions that serve as cleaning
CC solutions in a wide variety of applications including laundry detergents,
CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
CC represent the amino acid sequences of novel subtilisin homologues of the
CC invention.
XX
SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
| | | |
Db 71 gtsaps 76

Query Match 91.7%; Score 22; DB 22; Length 173;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
| | | |
Db 71 gtsaps 76

RESULT 41
AAU38780
ID AAU38780 standard; Protein; 173 AA.
XX AC AAU38780;
DT 19-DEC-2001 (first entry)
XX DE Novel subtilisin homologue #37.
XX KW Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
KW contact lens cleansing solution; dry cleaning.
XX OS Bacillus sp.
XX PN WO200175087-A2.
XX PD 11-OCT-2001.
XX PF 02-APR-2001; 2001WO-US10781.
XX PR 03-APR-2000; 2000US-194143P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
PI Minshull J;
XX WPI; 2001-616689/71.
DR N-PSDB; AAS56635.
XX Novel subtilisin homologue polypeptides having improved endo-protease
PT activity relative to mature Savinase subtilisin polypeptide, useful as
PT component of cleaning solutions e.g. laundry detergents, dry cleaning
PT

Claim 1; Page 106; 138pp; English.
The invention relates to an isolated polypeptide with improved endo-
protease activity relative to subtilisin homologue polypeptide, Savinase
(RTM). The invention also relates to an integrated system comprising a
computer or computer readable medium comprising a database comprising one
or more sequence records. Each record comprises one or more character
string corresponding to a nucleic acid or protein sequence of AAS56599-
AAS56728 or AAU38744-AAU38873 respectively. The integrated system
comprises a user input interface allowing a user to selectively
one or more sequence record. The integrated system is useful for
presenting information pertaining to one of several sequence records
stored in a database. The method involves determining a list of one or
more character strings corresponding to the sequence or its subsequence,
determining which character strings of the list are selected by a user,
and displaying the selected character strings or aligning the selected
character string with an additional character string. The method
further involves displaying an alignment of the selected character
string with the additional character string, and displaying the list.
The polynucleotides are useful as probes, primers, sense and antisense
nucleotides. The subtilisin homologue polypeptides are useful for
producing antibodies which have diagnostic uses related to the activity,
distribution and expression of subtilisin homologues. The subtilisin
homologue polypeptides are used in compositions that serve as cleaning
solutions in a wide variety of applications including laundry detergents,
contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
represent the amino acid sequences of novel subtilisin homologues of the
invention.

Sequence 173 AA;
XX SQ

QY 1 gtxxps 6
| | | |
Db 71 gtsaps 76

Query Match 91.7%; Score 22; DB 22; Length 173;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
| | | |
Db 71 gtsaps 76

RESULT 41
AAU38787
ID AAU38787 standard; Protein; 173 AA.
XX AC AAU38787;
DT 19-DEC-2001 (first entry)
XX DE Novel subtilisin homologue #44.
XX KW Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
KW contact lens cleansing solution; dry cleaning.
XX OS Bacillus sp.
XX PN WO200175087-A2.
XX PD 11-OCT-2001.
XX PF 02-APR-2001; 2001WO-US10781.
XX PR 03-APR-2000; 2000US-194143P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
PI Minshull J;
XX WPI; 2001-616689/71.
DR N-PSDB; AAS56642.
XX Novel subtilisin homologue polypeptides having improved endo-protease
PT activity relative to mature Savinase subtilisin polypeptide, useful as
PT component of cleaning solutions e.g. laundry detergents, dry cleaning
PT

Claim 1; Page 107; 138pp; English.
The invention relates to an isolated polypeptide with improved endo-
protease activity relative to subtilisin homologue polypeptide, Savinase
(RTM). The invention also relates to an integrated system comprising a
computer or computer readable medium comprising a database comprising one
or more sequence records. Each record comprises one or more character
string corresponding to a nucleic acid or protein sequence of AAS56599-
AAS56728 or AAU38744-AAU38873 respectively. The integrated system
comprises a user input interface allowing a user to selectively
one or more sequence record. The integrated system is useful for
presenting information pertaining to one of several sequence records
stored in a database. The method involves determining a list of one or
more character strings corresponding to the sequence or its subsequence,
determining which character strings of the list are selected by a user,
and displaying the selected character strings or aligning the selected
character string with an additional character string. The method
further involves displaying an alignment of the selected character
string with the additional character string, and displaying the list.
The polynucleotides are useful as probes, primers, sense and antisense
nucleotides. The subtilisin homologue polypeptides are useful for
producing antibodies which have diagnostic uses related to the activity,
distribution and expression of subtilisin homologues. The subtilisin
homologue polypeptides are used in compositions that serve as cleaning
solutions in a wide variety of applications including laundry detergents,
contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
represent the amino acid sequences of novel subtilisin homologues of the
invention.

CC invention.

XX Sequence 173 AA;

SQ Query Match 91.7%; Score 22; DB 22; Length 173;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 71 gtsps 76

RESULT 42
AAU38804
ID AAU38804 standard; Protein; 173 AA.

XX AC AAU38804;

XX DT 19-DEC-2001 (first entry)

XX Novel subtilisin homologue #61.

XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
DE contact lens cleansing solution; dry cleaning.

XX Bacillus sp.

XX WO200175087-A2.

XX 11-OCT-2001.

XX 02-APR-2001; 2001WO-US10781.

XX 03-APR-2000; 2000US-194143P.

XX (MAXY-) MAXYGEN INC.

XX Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
PI Minshull J;

XX WPI; 2001-616689/71.

XX N-PSDB; AAS56659.

XX Novel subtilisin homologue polypeptides having improved endo-protease
PT activity relative to mature Savinase subtilisin polypeptide, useful as
PT component of cleaning solutions e.g. laundry detergents, dry cleaning

PS Claim 1; Page 109; 139pp; English.

XX The invention relates to an isolated polypeptide with improved endo-
CC protease activity relative to subtilisin homologue polypeptide, Savinase
CC (RTM). The invention also relates to an integrated system comprising a
CC computer or computer readable medium comprising a database comprising one
CC or more sequence records. Each record comprises one or more character
CC string corresponding to a nucleic acid or protein sequence of AAS56599-
CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
CC comprises a user input interface allowing a user to selectively
CC presenting information pertaining to one of several sequence records
CC stored in a database. The method involves determining a list of one or
CC more character strings corresponding to the sequence or its subsequence,
CC and displaying the selected character strings or aligning the selected
CC character string with an additional character string. The method
CC further involves displaying an alignment of the selected character
CC string with the additional character string, and displaying the list.
CC The polynucleotides are useful as probes, primers, sense and antisense
CC nucleotides. The subtilisin homologue polypeptides are useful for
CC producing antibodies which have diagnostic uses related to the activity,
CC distribution and expression of subtilisin homologues. The subtilisin

CC homologue polypeptides are used in compositions that serve as cleaning
CC solutions in a wide variety of applications including laundry detergents,
CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
CC represent the amino acid sequences of novel subtilisin homologues of the
CC invention.

XX Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 71 gtsps 76

RESULT 43
AAU38810
ID AAU38810 standard; Protein; 173 AA.

XX AC AAU38810;

XX DT 19-DEC-2001 (first entry)

XX Novel subtilisin homologue #67.

XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
KW contact lens cleansing solution; dry cleaning.

XX Bacillus sp.

XX WO200175087-A2.

XX 11-OCT-2001.

XX 02-APR-2001; 2001WO-US10781.

XX 03-APR-2000; 2000US-194143P.

XX (MAXY-) MAXYGEN INC.

XX Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
PI Minshull J;

XX WPI; 2001-616689/71.

XX N-PSDB; AAS56665.

XX Novel subtilisin homologue polypeptides having improved endo-protease
PT activity relative to mature Savinase subtilisin polypeptide, useful as
PT component of cleaning solutions e.g. laundry detergents, dry cleaning

PS Claim 1; Page 110; 138pp; English.

XX The invention relates to an isolated polypeptide with improved endo-
CC protease activity relative to subtilisin homologue polypeptide, Savinase
CC (RTM). The invention also relates to an integrated system comprising a
CC computer or computer readable medium comprising a database comprising one
CC or more sequence records. Each record comprises one or more character
CC string corresponding to a nucleic acid or protein sequence of AAS56599-
CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
CC comprises a user input interface allowing a user to selectively
CC presenting information pertaining to one of several sequence records
CC stored in a database. The method involves determining a list of one or
CC more character strings corresponding to the sequence or its subsequence,
CC and displaying the selected character strings or aligning the selected
CC character string with an additional character string. The method
CC further involves displaying an alignment of the selected character
CC string with the additional character string, and displaying the list.

CC The polynucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

XX SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 71 gtsaps 76

RESULT 44
 AAU38835
 ID AAU38835 standard; Protein; 173 AA.

AC AAU38835;

XX 19-DEC-2001 (first entry)

XX Novel subtilisin homologue #92.

XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.

OS Bacillus sp.

XX WO200175087-A2.

XX 11-OCT-2001.

XX 02-APR-2001; 2001WO-US10781.

XX 03-APR-2000; 2000US-194143P.

XX (MAXY-) MAXYGEN INC.

XX Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;

DR WPI; 2001-616689/71.
 DR N-PSDB; AAS56690.

XX Novel subtilisin homologue polypeptides having improved endo-protease
 PT activity relative to mature Savinase subtilisin polypeptide, useful as
 PT component of cleaning solutions e.g. laundry detergents, dry cleaning
 PT

XX Claim 1; Page 113; 138pp; English.

XX The invention relates to an isolated polypeptide with improved endo-
 CC protease activity relative to subtilisin homologue polypeptide, Savinase
 CC (RTM). The invention also relates to an integrated system comprising a
 CC computer or computer readable medium comprising a database comprising one
 CC or more sequence records. Each record comprises one or more character
 CC string corresponding to a nucleic acid or protein sequence of AAS56599-
 CC AA556728 or AAU38744-AAU38873 respectively. The integrated system
 CC comprises a user input interface allowing a user to selectively
 CC one or more sequence record. The integrated system is useful for
 CC presenting information pertaining to one of several sequence records
 CC stored in a database. The method involves determining a list of one or
 CC more character strings corresponding to the sequence or its subsequence,
 CC determining which character strings of the list are selected by a user,

CC and displaying the selected character strings or aligning the selected
 CC character string with an additional character string. The method
 CC further involves displaying an alignment of the selected character
 CC string with the additional character string, and displaying the list.
 CC The polynucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

XX SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 71 gtsaps 76

RESULT 45

AAU38843

ID AAU38843 standard; Protein; 173 AA.

XX AC AAU38843;

XX 19-DEC-2001 (first entry)

XX Novel subtilisin homologue #100.

XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.

XX Bacillus sp.

XX WO200175087-A2.

XX 11-OCT-2001.

XX 02-APR-2001; 2001WO-US10781.

XX 03-APR-2000; 2000US-194143P.

XX (MAXY-) MAXYGEN INC.

XX Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;

DR WPI; 2001-616689/71.
 DR N-PSDB; AAS56698.

XX Novel subtilisin homologue polypeptides having improved endo-protease
 PT activity relative to mature Savinase subtilisin polypeptide, useful as
 PT component of cleaning solutions e.g. laundry detergents, dry cleaning
 PT

XX Claim 1; Page 114; 138pp; English.

XX The invention relates to an isolated polypeptide with improved endo-
 CC protease activity relative to subtilisin homologue polypeptide, Savinase
 CC (RTM). The invention also relates to an integrated system comprising a
 CC computer or computer readable medium comprising a database comprising one
 CC or more sequence records. Each record comprises one or more character
 CC string corresponding to a nucleic acid or protein sequence of AAS56599-
 CC AA556728 or AAU38744-AAU38873 respectively. The integrated system
 CC comprises a user input interface allowing a user to selectively
 CC one or more sequence record. The integrated system is useful for

CC presenting information pertaining to one of several sequence records
 CC stored in a database. The method involves determining a list of one or
 CC more character strings corresponding to the sequence or its subsequence,
 CC determining which character strings of the list are selected by a user,
 CC and displaying the selected character strings or aligning the selected
 CC character string with an additional character string. The method
 CC further involves displaying an alignment of the selected character
 CC string with the additional character string, and displaying the list.
 CC The polynucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

XX Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;

Best Local Similarity 66.7%; Pred. No. 6.3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 71 gtspps 76

RESULT 46

AAU38869

ID AAU38869 standard; Protein; 173 AA.

XX

AC AAU38869;

XX

DT 19-DEC-2001 (first entry)

XX

DE Novel subtilisin homologue #126.

XX

KW Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.

XX

OS Bacillus sp.

XX

PN WO200175087-A2.

XX

PD 11-OCT-2001.

XX

PF 02-APR-2001; 2001WO-US10781.

XX

PR 03-APR-2000; 2000US-194143P.

XX

PA (MAXY-) MAXYGEN INC.

XX

PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;

XX

DR WPI; 2001-616689/71.

XX

DR N-PSDB; AAS56724.

XX

PT Novel subtilisin homologue polypeptides having improved endo-protease
 PT activity relative to mature Savinase subtilisin polypeptide, useful as
 PT component of cleaning solutions e.g. laundry detergents, dry cleaning

XX

PS Claim 1; Page 117; 138pp; English.

XX

CC The invention relates to an isolated polypeptide with improved endo-
 CC protease activity relative to subtilisin homologue polypeptide, Savinase
 CC (RTM). The invention also relates to an integrated system comprising a
 CC computer or computer readable medium comprising a database comprising one
 CC or more sequence records. Each record comprises one or more character

CC string corresponding to a nucleic acid or protein sequence of AAS56599-
 CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 CC comprises a user input interface allowing a user to selectively
 CC one or more sequence record. The integrated system is useful for
 CC presenting information pertaining to one of several sequence records
 CC stored in a database. The method involves determining a list of one or
 CC more character strings corresponding to the sequence or its subsequence,
 CC determining which character strings of the list are selected by a user,
 CC and displaying the selected character strings or aligning the selected
 CC character string with an additional character string. The method
 CC further involves displaying an alignment of the selected character
 CC string with the additional character string, and displaying the list.
 CC The polynucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;

Best Local Similarity 66.7%; Pred. No. 6.3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 71 gtspps 76

RESULT 47

ABB68003

ID ABB68003 standard; Protein; 175 AA.

XX

AC ABB68003;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 30801.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

XX

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

XX

DR N-PSDB; ABL12106.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX

PS Disclosure; SEQ ID NO 30801; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 175 AA;

Query Match 91.7%; Score 22; DB 22; Length 175;
Best Local Similarity 66.7%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 98 gdstps 103

RESULT 48

ABB68002
ID ABB68002 standard; Protein; 182 AA.

AC ABB68002;

XX
XX 26-MAR-2002 (first entry)

XX
XX Drosophila melanogaster polypeptide SEQ ID NO 30798.

XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX
XX Drosophila melanogaster.

XX
XX W0200171042-A2.

XX
XX 27-SEP-2001.

XX
XX 23-MAR-2001; 2001WO-US09231.

XX
XX 23-MAR-2000; 2000US-191637P.

XX
XX 11-JUL-2000; 2000US-0614150.

XX
XX (PEKE) PE CORP NY.

XX
XX Venter JC, Adams M, Li PWD, Myers EW;

XX
XX WPI; 2001-6556860/75.

XX
XX N-PSDB; ABL12105.

XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX
XX Disclosure; SEQ ID NO 30798; 2hpp + Sequence Listing; English.

XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).

XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX
XX Sequence 182 AA;

Query Match 91.7%; Score 22; DB 22; Length 182;
Best Local Similarity 66.7%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 105 gdstps 110

RESULT 49

AAM40840

ID AAM40840 standard; Protein; 190 AA.

XX
XX AAM40840;

XX
XX 22-OCT-2001 (first entry)

XX
XX Human polypeptide SEQ ID NO 5771.

XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX
XX Homo sapiens.

XX
XX W0200153312-A1.

XX
XX 26-JUL-2001.

XX
XX 26-DEC-2000; 2000WO-US34263.

XX
XX 21-JAN-2000; 2000US-0488725.

XX
XX 25-APR-2000; 2000US-0552317.

XX
XX 09-JUL-2000; 2000US-0598042.

XX
XX 19-JUL-2000; 2000US-0620312.

XX
XX 03-AUG-2000; 2000US-0653450.

XX
XX 14-SEP-2000; 2000US-0662191.

XX
XX 19-OCT-2000; 2000US-0693036.

XX
XX 29-NOV-2000; 2000US-0727344.

XX
XX (HYSE-) HYSEQ INC.

XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX
XX WPI; 2001-442253/47.

XX
XX N-PSDB; AAI59996.

XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -

XX
XX Example 2; SEQ ID NO 5771; 10078pp; English.

XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities, such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.

XX
XX Note: The sequence data for this patent did not form part of the printed

CC specification.
XX
SQ Sequence 190 AA;

Query Match 91.7%; Score 22; DB 22; Length 190;
Best Local Similarity 66.7%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
| | | |
Db 39 gtstps 44

RESULT 50
AAU40065
ID AAU40065 standard; Protein; 194 AA.
AC AAU40065;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #961.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59510.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 1260; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC
CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 194 AA;

Query Match 91.7%; Score 22; DB 22; Length 194;
Best Local Similarity 66.7%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
| | | |
Db 48 gtssps 53

Search completed: September 24, 2002, 11:22:19
Job time: 6018 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:23:07 ; Search time 36.48 Seconds
(without alignments)
4.017 Million cell updates/sec

Title: BASK-853-CLAIM5
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	38	1	US-08-176-500-41
2	22	91.7	38	1	US-08-471-052A-41
3	22	91.7	38	1	US-08-189-331-41
4	22	91.7	38	2	US-08-471-939-41
5	22	91.7	38	2	US-08-471-800-41
6	22	91.7	38	2	US-08-471-068-41
7	22	91.7	166	1	US-08-665-617-2
8	22	91.7	168	2	US-08-353-476-73
9	22	91.7	241	2	US-08-353-476-116
10	22	91.7	273	2	US-08-353-476-112
11	22	91.7	294	4	US-09-372-422A-40
12	22	91.7	1050	4	US-09-428-711A-16
13	21	87.5	40	5	PCT-US96-08730-14
14	21	87.5	56	2	US-08-537-400-32
15	21	87.5	153	4	US-09-228-986-83
16	21	87.5	170	2	US-08-483-101-3
17	21	87.5	170	3	US-09-081-180-5
18	21	87.5	170	3	US-09-040-786-5
19	21	87.5	225	1	US-08-290-979A-8
20	21	87.5	295	2	US-08-481-956A-9
21	21	87.5	295	2	US-08-629-291A-9
22	21	87.5	295	2	US-08-658-335B-9
23	21	87.5	334	2	US-08-997-080-45
24	21	87.5	334	2	US-08-997-362-45
25	21	87.5	334	3	US-08-873-970-45
26	21	87.5	334	4	US-09-095-855-45
27	21	87.5	334	4	US-08-705-347A-45

ALIGNMENTS

RESULT 1
US-08-176-500-41
; Sequence 41, Application US/08176500
; Patent No. 5498538
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-176-500-41

Sequence 45, Appl
Sequence 10, Appl
Sequence 1, Appl
Sequence 184, App
Sequence 2, Appl
Sequence 3, Appl
Sequence 32, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 28, Appl
Sequence 2, Appl
Sequence 42, Appl
Sequence 42, Appl
Sequence 42, Appl

28 21 87.5 334 4 US-09-324-542-45
29 21 87.5 335 4 US-09-198-955A-10
30 21 87.5 335 4 US-09-184-217-1
31 21 87.5 340 4 US-09-188-930-184
32 21 87.5 385 3 US-09-053-866-2
33 21 87.5 389 1 US-08-409-199-3
34 21 87.5 439 4 US-09-171-461-32
35 21 87.5 446 3 US-08-956-254-2
36 21 87.5 446 3 US-09-008-388-1
37 21 87.5 448 2 US-09-015-815-1
38 21 87.5 448 3 US-08-476-509B-2
39 21 87.5 486 3 US-08-348-518C-2
40 21 87.5 495 2 US-08-794-795-2
41 21 87.5 495 4 US-09-249-200-2
42 21 87.5 534 2 US-09-031-392-4
43 21 87.5 534 4 US-09-299-549-4
44 21 87.5 534 4 US-09-610-417-4
45 21 87.5 577 4 US-09-413-814-4
46 21 87.5 732 2 US-08-843-530B-28
47 21 87.5 739 2 US-08-836-943-2
48 21 87.5 774 2 US-08-231-193A-42
49 21 87.5 774 2 US-08-486-273A-42
50 21 87.5 774 3 US-08-480-474-42

Query Match 91.7%; Score 22; DB 1; Length 38;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 20 GTTSPS 25

RESULT 2
US-08-471-052A-41
; Sequence 41, Application US/08471052A
; Patent No. 5625033
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471.052A
; FILING DATE: 06-JUNE-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-471-052A-41

Query Match 91.7%; Score 22; DB 1; Length 38;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 20 GTTSPS 25

RESULT 3
US-08-189-331-41
; Sequence 41, Application US/08189331
; Patent No. 5747334
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 186
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/189,331
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-155
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-189-331-41

Query Match 91.7%; Score 22; DB 1; Length 38;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 20 GTTSPS 25

RESULT 4
US-08-471-939-41
; Sequence 41, Application US/08471939
; Patent No. 5844076
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,939
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,416
; FILING DATE: 01-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-471-939-41

Query Match 91.7%; Score 22; DB 2; Length 38;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 20 GTTSPS 25

RESULT 5
US-08-471-800-41
; Sequence 41, Application US/08471800
; Patent No. 5852167
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,800
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/013,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-471-800-41

Query Match 91.7%; Score 22; DB 2; Length 38;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 20 GTTSPS 25

RESULT 6
US-08-471-068-41
; Sequence 41, Application US/08471068
; Patent No. 5948635
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 186
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,068
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/189,331
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-155
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-471-068-41

Query Match 91.7%; Score 22; DB 2; Length 38;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 20 GTTSPS 25

RESULT 7
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
; GENERAL INFORMATION:
; APPLICANT: Xudong, Yin
; TITLE OF INVENTION: Gene and Protein for Regulation of Cell Death
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida

; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,617
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: CL-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-665-617-2

Query Match 91.7%; Score 22; DB 1; Length 166;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 160 GTAAPS 165

RESULT 8
US-08-353-476-73
; Sequence 73, Application US/08353476
; Patent No. 5871902
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,476
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-353-476-73

Query Match 91.7%; Score 22; DB 2; Length 168;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 119 GTATPS 124

RESULT 9
US-08-353-476-116
; Sequence 116, Application US/08353476
; Patent No. 5871902
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,476
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-353-476-116

Query Match 91.7%; Score 22; DB 2; Length 241;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 192 GTATPS 197

RESULT 10

US-08-353-476-112
; Sequence 112, Application US/08353476
; Patent No. 5871902
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/353,476
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-353-476-112

Query Match 91.7%; Score 22; DB 2; Length 273;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 224 GTATPS 229

RESULT 11
US-09-372-422A-40
; Sequence 40, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Zea mays
US-09-372-422A-40

Query Match 91.7%; Score 22; DB 4; Length 294;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 158 GTTTPS 163

RESULT 12
US-09-428-711A-16
; Sequence 16, Application US/09428711A
; Patent No. 6358720
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masaaki
; APPLICANT: Shirasawa, Takuji
; APPLICANT: Tokumitsu, Hiroshi
; APPLICANT: No. 6358720uchl, Teruhisa
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
; FILE REFERENCE: 06501-045001
; CURRENT APPLICATION NUMBER: US/09/428,711A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: PCT/JP98/01246
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: JP 9/124798
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1050
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-428-711A-16

Query Match 91.7%; Score 22; DB 4; Length 1050;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 745 GTSSPS 750

RESULT 13
PCT-US96-08730-14
; Sequence 14, Application PC/TUS9608730
; GENERAL INFORMATION:
; APPLICANT: Cassels, Frederick
; APPLICANT: Anderson, Jeffrey
; APPLICANT: Carter, John Mark
; TITLE OF INVENTION: Methods of Raising Antibodies Against E.
; TITLE OF INVENTION: Coli of the Family CSF-CFA./1
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glenna Hendricks
; STREET: P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: USA
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08730
; FILING DATE: 03-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: PCT/US96/08730
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 591-4470
TELEFAX: (703) 591-4428
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
PCT-US96-08730-14

Query Match 87.5%; Score 21; DB 5; Length 40;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 22 GTALPS 27

RESULT 14
US-08-537-400-32
; Sequence 32, Application US/08537400
; Patent No. 5939301
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cloned DNA Polymerases From Thermotoga
; TITLE OF INVENTION: neapolitana And Mutants Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,400
; FILING DATE: 02-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,423
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/370,190
; FILING DATE: 09-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-537-400-32

Query Match 87.5%; Score 21; DB 2; Length 56;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 44 GTRTPS 49

RESULT 15
US-09-228-986-83
; Sequence 83, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 83
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-83

Query Match 87.5%; Score 21; DB 4; Length 153;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 103 GTLSPS 108

RESULT 16
US-08-483-101-3
; Sequence 3, Application US/08483101
; Patent No. 5932715
; GENERAL INFORMATION:
; APPLICANT: Scott, June R.
; APPLICANT: Froehlich, Barbara
; APPLICANT: Caron, Judy
; TITLE OF INVENTION: CS2 Proteins and Coding Sequences
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,101
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33878
; REFERENCE/DOCKET NUMBER: 6-95
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-483-101-3

Query Match 87.5%; Score 21; DB 2; Length 170;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 qtxxps 6
|| ||
Db 45 GTALPS 50

RESULT 17
US-09-081-180-5
; Sequence 5, Application US/09081130
; Patent No. 6022847
; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG32
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics
; STREET: 1201 Eastlake Ave. E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,180
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,263
; FILING DATE: March 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan E
; REGISTRATION NUMBER: 41,156
; REFERENCE/DOCKET NUMBER: 97-17C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-180-5

Query Match 87.5%; Score 21; DB 3; Length 170;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 qtxxps 6
|| ||

Db 95 GTYSPS 100
RESULT 18
US-09-040-786-5
; Sequence 5, Application US/09040786
; Patent No. 6025197
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG32
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics
; STREET: 1201 Eastlake Ave. E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,786
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,263
; FILING DATE: March 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan E
; REGISTRATION NUMBER: 41,156
; REFERENCE/DOCKET NUMBER: 97-17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-040-786-5

Query Match 87.5%; Score 21; DB 3; Length 170;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 qtxxps 6
|| ||
Db 95 GTYSPS 100

RESULT 19
US-08-290-979A-8
; Sequence 8, Application US/08290979A
; Patent No. 5610046
; GENERAL INFORMATION:

; APPLICANT: VAN OIJEN, Albert J.H.
; APPLICANT: DE GRAAFF, Leendert H.
; APPLICANT: VAN DEN BROECK, Henriette C.
; APPLICANT: VISSER, Jacob
; TITLE OF INVENTION: Cloning and Expression of Xylanase B
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,979A
; FILING DATE: 22-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KATE H. MURASHIGE
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 4615-0045.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030 MRSN FOERS WSH
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-290-979A-8

Query Match 87.5%; Score 21; DB 1; Length 225;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 99 GTTPS 104

RESULT 20
US-08-481-956A-9
; Sequence 9, Application US/08481956A
; Patent No. 5824867
; GENERAL INFORMATION:
; APPLICANT: Coruzzi, Gloria
; APPLICANT: Oliveira, Igor
; APPLICANT: Lam, Hon-Ming
; APPLICANT: Hsieh, Ming-Hsiun
; TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,956A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 5914-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-481-956A-9

Query Match 87.5%; Score 21; DB 2; Length 295;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 65 GTVSPS 70

RESULT 21
US-08-629-291A-9
; Sequence 9, Application US/08629291A
; Patent No. 5959174
; GENERAL INFORMATION:
; APPLICANT: Coruzzi, Gloria
; APPLICANT: Oliveira, Igor
; APPLICANT: Lam, Hon-Ming
; APPLICANT: Hsieh, Ming-Hsiun
; TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/629,291A
; FILING DATE: 08-APR-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 5914-050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-629-291A-9

Query Match 87.5%; Score 21; DB 2; Length 295;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 65 GTVSPS 70

```
RESULT 22
US-08-658-335B-9
; Sequence 9, Application US/08658335B
; Patent No. 5981703
; GENERAL INFORMATION:
; APPLICANT: Coruzzi, Gloria
; APPLICANT: Oliveira, Igor
; APPLICANT: Lam, Hon-Ming
; APPLICANT: Hsieh, Ming-Hsiun
; TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,335B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 5914-052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-658-335B-9

Query Match 87.5%; Score 21; DB 2; Length 295;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   |||
Db 65 GTVSPS 70

RESULT 23
US-08-997-080-45
; Sequence 45, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-080-45

Query Match 87.5%; Score 21; DB 2; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   |||
Db 253 GTGTPS 258

RESULT 24
US-08-997-362-45
; Sequence 45, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
```

; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-362-45

Query Match 87.5%; Score 21; DB 2; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 253 GTGTPS 258

RESULT 25
US-08-873-970-45
; Sequence 45, Application US/08873970
; Patent No. 6001361
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-873-970-45

Query Match 87.5%; Score 21; DB 3; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 253 GTGTPS 258

RESULT 26
US-09-095-855-45
; Sequence 45, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TREATMENT AND DIAGNOSIS OF Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-095-855-45

Query Match 87.5%; Score 21; DB 4; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
|| ||
Db 253 GTGTPS 258


```
RESULT 27
US-08-705-347A-45
; Sequence 45, Application US/08705347A
; Patent No. 6284255
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,347A
; FILING DATE: 28-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206.269.0565
; TELEFAX: 206.269.0563
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-705-347A-45

Query Match 87.5%; Score 21; DB 4; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 253 GTGTPS 258

RESULT 28
US-09-324-542-45
; Sequence 45, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-324-542-45

Query Match 87.5%; Score 21; DB 4; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 253 GTGTPS 258

RESULT 29
US-09-198-955A-10
; Sequence 10, Application US/09198955A
; Patent No. 6187580
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6187580el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/09/198,955A
; CURRENT FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-198-955A-10

Query Match 87.5%; Score 21; DB 4; Length 335;
Best Local Similarity 66.7%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 84 GTITPS 89

RESULT 30
US-09-184-217-1
; Sequence 1, Application US/09184217
; Patent No. 6258590
; GENERAL INFORMATION:
; APPLICANT: Lange, Niels E.K.
; APPLICANT: Kongsbak, Lars
; APPLICANT: Schulein, Martin
```

; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Husain, Philip A.
; TITLE OF INVENTION: Biopreparation Of Textiles At High
; FILE REFERENCE: 5729.000-US
; CURRENT APPLICATION NUMBER: US/09/184,217
; CURRENT FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: bacillus sp.
US-09-184-217-1

Query Match 87.5%; Score 21; DB 4; Length 335;
Best Local Similarity 66.7%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 84 GTITPS 89

RESULT 31
US-09-188-930-184
; Sequence 184, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 340
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-184

Query Match 87.5%; Score 21; DB 4; Length 340;
Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 120 GTAGPS 125

RESULT 32
US-09-053-866-2
; Sequence 2, Application US/09053866
; Patent No. 6111075
; GENERAL INFORMATION:
; APPLICANT: Xu, Wenfeng
; APPLICANT: Presnell, Scott R.
; APPLICANT: Yee, David P.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: PROTEASE-ACTIVATED RECEPTOR
; FILE REFERENCE: PAR4 (ZCHEMR2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East

; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leith, Debra K
; REGISTRATION NUMBER: 32,619
; REFERENCE/DOCKET NUMBER: 98-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-09-053-866-2

Query Match 87.5%; Score 21; DB 3; Length 385;
Best Local Similarity 66.7%; Pred. No. 9.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 19 GTQTPS 24

RESULT 33
US-08-409-199-3
; Sequence 3, Application US/08409199
; Patent No. 5532153
; GENERAL INFORMATION:
; APPLICANT: XU, Shuang-Yong
; APPLICANT: XIAO, Jianping
; TITLE OF INVENTION: METHOD FOR CLONING AND
; TITLE OF INVENTION: PRODUCING THE SacI RESTRICTION
; TITLE OF INVENTION: ENDONUCLEASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEW ENGLAND BIOLABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/409,199
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.

; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054; 292
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-409-199-3

Query Match 87.5%; Score 21; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 375 GTTIPS 380

RESULT 34

US-09-171-461-32
; Sequence 32, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schafner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 439
; TYPE: PRT
; ORGANISM: CELO VIRUS
; FEATURE:
; OTHER INFORMATION: Position: 5366..6685/Product: IVa2
US-09-171-461-32

Query Match 87.5%; Score 21; DB 4; Length 439;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 19 GTKTPS 24

RESULT 35

US-08-956-254-2
; Sequence 2, Application US/08956254A
; Patent No. 6013265
; GENERAL INFORMATION:
; APPLICANT: AURELIAN, LAURE
; TITLE OF INVENTION: Vaccine Composition for Herpes Simplex Virus and
; FILE REFERENCE: 14211A
; CURRENT APPLICATION NUMBER: US/08/956,254A
; CURRENT FILING DATE: 1997-10-22
; EARLIER APPLICATION NUMBER: US 60/029,093
; EARLIER FILING DATE: 1996-10-22
; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRT
; ORGANISM: herpes simplex virus-2
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(446)
; OTHER INFORMATION: PROTEIN KINASE DOMAIN OF ICP10 SUBUNIT OF HSV-2
; PUBLICATION INFORMATION:
; AUTHORS: CHUNG ET AL.,
; JOURNAL: J. Virol.
; VOLUME: 63
; PAGES: 3389-3398
; DATE: 1989
; PUBLICATION INFORMATION:
; AUTHORS: NELSON ET AL.,
; JOURNAL: J. Biol. Chem.
; VOLUME: 271
; PAGES: 17021-17027
; DATE: 1996
US-08-956-254-2

Query Match 87.5%; Score 21; DB 3; Length 446;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 116 GTSGPS 121

RESULT 36

US-09-008-388-1
; Sequence 1, Application US/09008388
; Patent No. 6054131
; GENERAL INFORMATION:
; APPLICANT: AURELIAN, LAURE
; TITLE OF INVENTION: VACCINE COMPOSITION FOR HERPES
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PEPPER HAMILTON LLP
; STREET: 600 FOURTEENTH STREET, N.W., Suite 500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005-2004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-WINDOWS
; SOFTWARE: WP 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,388
; FILING DATE: December 2, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramsey, William S.
; REGISTRATION NUMBER: 32,715
; REFERENCE/DOCKET NUMBER: 114178.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 220-1280
; TELEFAX: (202) 220-1665
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE:
ORGANISM: HERPES SIMPLEX
STRAIN: RECOMBINANT
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-008-388-1

Query Match 87.5%; Score 21; DB 3; Length 446;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 116 GTSGPS 121

RESULT 37

US-09-015-815-1
; Sequence 1, Application US/09015815
; Patent No. 5965356
; GENERAL INFORMATION:
; APPLICANT: AURELIAN, LAURE
; APPLICANT: SMITH, CYNTHIA
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS TYPE SPECIFIC SEROASSAY
; FILE REFERENCE: 1437LA
; CURRENT APPLICATION NUMBER: US/09/015,815
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: US 60/036,622
; EARLIER FILING DATE: 1997-01-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 448
; TYPE: PRT
; ORGANISM: herpes simplex virus-2
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(446)
; OTHER INFORMATION: PROTEIN KINASE DOMAIN OF ICP10 SUBUNIT OF HSV-2
; PUBLICATION INFORMATION:
; AUTHORS: CHUNG ET AL.,
; JOURNAL: J. Virol.
; VOLUME: 63
; PAGES: 3389-3398
; DATE: 1989
; PUBLICATION INFORMATION:
; AUTHORS: NELSON ET AL.,
; JOURNAL: J. Biol. Chem.
; VOLUME: 271
; PAGES: 17021-17027
; DATE: 1996
US-09-015-815-1

Query Match 87.5%; Score 21; DB 2; Length 448;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
|| ||
Db 116 GTSGPS 121
RESULT 38
US-08-476-509B-2
; Sequence 2, Application US/08476509B
; Patent No. 6034212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTICUSES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-509B-2

Query Match 87.5%; Score 21; DB 3; Length 448;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 139 GTLTFS 144

RESULT 39

US-08-348-518C-2
; Sequence 2, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A

```
; TITLE OF INVENTION:  SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION:  PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION:  THEREOF
; NUMBER OF SEQUENCES:  26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Klauber & Jackson
; STREET:  411 Hackensack Avenue
; CITY:  Hackensack
; STATE:  New Jersey
; COUNTRY:  USA
; ZIP:  07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/348,518C
; FILING DATE:  01-DEC-1994
; CLASSIFICATION:  514
; ATTORNEY/AGENT INFORMATION:
; NAME:  Jackson Esq., David A.
; REGISTRATION NUMBER:  26,742
; REFERENCE/DOCKET NUMBER:  600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  201 487-5800
; TELEFAX:  201 343-1684
; TELEX:  133521
; INFORMATION FOR SEQ ID NO:  2:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  486 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
; US-08-348-518C-2

Query Match      87.5%; Score 21; DB 3; Length 486;
Best Local Similarity 86.7%; Pred. No. 1.2e+03;
Matches  4; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

QY  1 gtxxps 6
   || ||
Db  139 GTLTPS 144

RESULT 40
US-08-794-795-2
; Sequence 2, Application US/08794795
; Patent No. 5916766
; GENERAL INFORMATION:
; APPLICANT:  Elshourlagy, Nabil
; APPLICANT:  Adamou, John
; APPLICANT:  Gross, Mitchell
; APPLICANT:  Lyoko, Paul
; TITLE OF INVENTION:  Human Macro Scavenger Rec
; TITLE OF INVENTION:  eptor
; NUMBER OF SEQUENCES:  9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  SmithKline Beecham Corporation
; STREET:  709 Swedeland Road
; CITY:  King of Prussia
; STATE:  PA
; COUNTRY:  USA
; ZIP:  19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Diskette
; COMPUTER:  IBM Compatible
; OPERATING SYSTEM:  DOS
; SOFTWARE:  FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/794,795
; FILING DATE:  04-FEB-1997
```

```
; CLASSIFICATION:  435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  ATG50009P
; FILING DATE:  22-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME:  Han, William T
; REGISTRATION NUMBER:  34,344
; REFERENCE/DOCKET NUMBER:  ATG50009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  610-270-5219
; TELEFAX:  610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO:  2:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  495 amino acids
; TYPE:  amino acid
; STRANDEDNESS:  single
; TOPOLOGY:  linear
; MOLECULE TYPE:  peptide
; US-08-794-795-2

Query Match      87.5%; Score 21; DB 2; Length 495;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches  4; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

QY  1 gtxxps 6
   || ||
Db  162 GTTGPS 167

RESULT 41
US-09-249-200-2
; Sequence 2, Application US/09249200
; Patent No. 6197931
; GENERAL INFORMATION:
; APPLICANT:  ELSHOURBAGY, NABIL
; APPLICANT:  ADAMOU, JOHN
; APPLICANT:  GROSS, MITCHELL
; APPLICANT:  LYSKO, PAUL
; TITLE OF INVENTION:  HUMAN MARCO SCAVENGER RECEPTOR
; NUMBER OF SEQUENCES:  9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Ratner & Prestia
; STREET:  P.O. Box 980
; CITY:  Valley Forge
; STATE:  PA
; COUNTRY:  USA
; ZIP:  19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Diskette
; COMPUTER:  IBM Compatible
; OPERATING SYSTEM:  DOS
; SOFTWARE:  FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/09/249,200
; FILING DATE:  12-FEB-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  08/794,795
; FILING DATE:  04-FEB-1997
; APPLICATION NUMBER:  60/017,699
; FILING DATE:  23-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME:  Prestia, Paul F
; REGISTRATION NUMBER:  23,031
; REFERENCE/DOCKET NUMBER:  ATG-50009-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  610-407-0700
; TELEFAX:  610-407-0700
; TELEX:  846169
; INFORMATION FOR SEQ ID NO:  2:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 495 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-249-200-2

```

```

Query Match      87.5%; Score 21; DB 4; Length 495;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 gtxxps 6
   ||||
Db 162 GTTGPS 167

```

```

RESULT 42
US-09-031-392-4
; Sequence 4, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-4

```

```

Query Match      87.5%; Score 21; DB 2; Length 534;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 gtxxps 6
   ||||
Db 87 GTLAPS 92

```

```

RESULT 43
US-09-299-549-4
; Sequence 4, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:

```

```

; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-549-4

```

```

Query Match      87.5%; Score 21; DB 4; Length 534;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 gtxxps 6
   ||||
Db 87 GTLAPS 92

```

```

RESULT 44
US-09-610-417-4
; Sequence 4, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417

```

```

; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-610-417-4

Query Match 87.5%; Score 21; DB 4; Length 534;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 87 GTLAPS 92

RESULT 45
US-09-413-814-4
; Sequence 4, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; EARLIER FILING DATE: 1999-10-07
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-4

Query Match 87.5%; Score 21; DB 4; Length 577;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 371 GTLAPS 376

RESULT 46
```

```

US-08-843-530B-28
; Sequence 28, Application US/08843530B
; Patent No. 5939306
; GENERAL INFORMATION:
; APPLICANT: Selitrennikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,530B
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Macknight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UTC-02717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-843-530B-28

Query Match 87.5%; Score 21; DB 2; Length 732;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 81 GTLAPS 86

RESULT 47
US-08-836-943-2
; Sequence 2, Application US/08836943
; Patent No. 5965391
; GENERAL INFORMATION:
; APPLICANT: Reinscheid, Dieter
; APPLICANT: Eikmanns, Bernhard
; APPLICANT: Sahm, Hermann
; TITLE OF INVENTION: DNA WHICH REGULATES GENE EXPRESSION IN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Firm of Karl F. Ross, PC
; STREET: 5676 Riverdale Ave.
; CITY: Bronx
; STATE: New York
; COUNTRY: USA
; ZIP: 10471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,943
; FILING DATE: 08-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Jonathan
; REGISTRATION NUMBER: 26,963
; REFERENCE/DOCKET NUMBER: 20357
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (718) 884-6600
; TELEFAX: 718/601-1099
; TELEX: 620428
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-836-943-2

Query Match 87.5%; Score 21; DB 2; Length 739;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 716 GTRKSPS 721

RESULT 48
US-08-231-193A-42
; Sequence 42, Application US/08231193A
; Patent No. 5849895
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,193A
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,459
; FILING DATE: 20-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-836-943-2

; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-193A-42

Query Match 87.5%; Score 21; DB 2; Length 774;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 292 GTVSPS 297

RESULT 49
US-08-486-273A-42
; Sequence 42, Application US/08486273A
; Patent No. 5985586
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; TITLE OF INVENTION: ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-273A-42

Query Match 87.5%; Score 21; DB 2; Length 774;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 292 GTVSPS 297

RESULT 50
US-08-480-474-42
; Sequence 42, Application US/08480474
; Patent No. 6033865
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA ENCODING
; TITLE OF INVENTION: SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,474
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,775
; REFERENCE/DOCKET NUMBER: 6362-9382B
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-474-42

Query Match 87.5%; Score 21; DB 3; Length 774;
Best Local Similarity 66.7%; Fred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 qtxxps 6
Db 292 GTVSPS 297

Search completed: September 24, 2002, 11:23:09
Job time: 4458 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright.(c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:24:17 ; Search time 54.84 Seconds
(without alignments)
10.513 Million cell updates/sec

Title: BASK-853-CLAIM5
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	91.7	89	E83141	hypothetical prote
2	22	91.7	101	S30493	Spl protein - mous
3	22	91.7	105	G72515	hypothetical prote
4	22	91.7	111	A85695	hypothetical prote
5	22	91.7	135	JC2424	hypothetical 14.3K
6	22	91.7	134	H50836	hypothetical prote
7	22	91.7	196	I39898	blue copper-bindin
8	22	91.7	196	T51838	probable outer mem
9	22	91.7	202	B82133	probable thioester
10	22	91.7	254	A83345	nitrite hydratase
11	22	91.7	255	E42725	myoD protein - zeb
12	22	91.7	275	S47325	hypothetical prote
13	22	91.7	280	H71320	hypothetical prote
14	22	91.7	285	H85356	hypothetical prote
15	22	91.7	289	T17957	hypothetical prote
16	22	91.7	316	T34553	hypothetical prote
17	22	91.7	319	S62196	hypothetical prote
18	22	91.7	379	JN0013	synaptic vesicle m
19	22	91.7	381	SL2223	naringenin-chalcon
20	22	91.7	383	E87549	type IV secretion
21	22	91.7	389	SVPCJ	naringenin-chalcon
22	22	91.7	389	JC5136	naringenin-chalcon
23	22	91.7	398	S42523	naringenin-chalcon
24	22	91.7	410	SL2224	naringenin-chalcon
25	22	91.7	418	F75587	probable glycosylt
26	22	91.7	419	SVPCJ	naringenin-chalcon
27	22	91.7	419	T36272	hypothetical prote
28	22	91.7	421	T41156	probable WD repeat
29	22	91.7	435	T24477	hypothetical prote

ALIGNMENTS

RESULT 1

E83141
hypothetical protein PA4033 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83141
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: E83141
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <STO>
A:Cross-references: GB:AE004820; GB:AE004091; NID:g9950223; PIDN:AAG07420.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4033

Query Match 91.7% Score 22; DB 2; Length 89;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 48 GTASPS 53

RESULT 2

S30493
Spl protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S30493
R:Chestier, A.; Charnay, P.
DNA Seq. 2, 325-327, 1992
A:Title: Difference in the genomic organizations of the related transcription factors
A:Reference number: S30493; MUID:92338398
A:Accession: S30493
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-101 <CHE>
A:Cross-references: EMBL:X60136; NID:g54158; PIDN:CAA42721.1; PID:e38120; PID:g133426

Query Match 91.7%; Score 22; DB 2; Length 101;
 Best Local Similarity 66.7%; Pred. No. 81;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 52 GTATPS 57

RESULT 3

G72515
 hypothetical protein APE2100 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: G72515
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339

A:Accession: G72515

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <RAW>

A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAA81111.1; PID:d1044897; PID:g510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2100

Query Match 91.7%; Score 22; DB 2; Length 105;
 Best Local Similarity 66.7%; Pred. No. 84;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 15 GTSSPS 20

RESULT 4

A85695
 hypothetical protein Z1932 [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: A85695
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew illier, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: A85695

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-111 <STO>

A:Cross-references: GB:AE005174; NID:g12514864; PIDN:AAC56021.1; GSPDB:GN00145; UWGP:219

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z1932

Query Match 91.7%; Score 22; DB 2; Length 111;
 Best Local Similarity 66.7%; Pred. No. 89;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 55 GTSSPS 60

RESULT 5

JC2424

hypothetical 14.3K protein - mouse

C:Species: Mus musculus (house mouse)
 C>Date: 01-Feb-1995 #sequence_revision 26-May-1995 #text_change 26-May-1995
 C:Accession: JC2424
 R:Harada, T.; Saito, H.; Kouhara, H.; Kurebayashi, S.; Kasayama, S.; Terakawa, N.; Ki Biochem. Biophys. Res. Commun. 205, 1057-1063, 1994
 A:Title: Murine fibroblast growth factor receptor 1 gene generates multiple messenger
 A:Reference number: PC2277; MUID:95100926
 A:Accession: JC2424
 A:Molecule type: DNA
 A:Residues: 1-135 <HAR>
 A:Note: The authors translated the codon CAG for residue 113 as Glu and AAC for resid

Query Match 91.7%; Score 22; DB 2; Length 135;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 105 GTAAPS 110

RESULT 6

H90836

hypothetical protein ECs1664 [imported] - Escherichia coli (strain O157:H7, substrain

C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: H90836
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
 A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H90836

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-194 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA035087.1; PID:g13361128; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECs1664

Query Match 91.7%; Score 22; DB 2; Length 194;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 49 GTSSPS 54

RESULT 7

I39698

blue copper-binding protein, 20K - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 11-Jun-1999

C:Accession: I39698; S25555

R:Van Gysel, A.; Van Montagu, M.; Inze, D. Gene 136, 79-85, 1993

A:Title: A negatively light-regulated gene from Arabidopsis thaliana encodes a protei

A:Reference number: I39698; MUID:94124044

A:Accession: I39698

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-196 <RES>

A:Cross-references: EMBL:Z15058; NID:g16202; PIDN:CAA78771.1; PID:g16203

C:Genetics:

A:Gene: bcb

A:Introns: 57/3

C:Superfamily: plastocyanin

C:Keywords: copper binding; transmembrane protein

F:66,107,112,117/Binding site: copper (His, Cys, His, Gln) #status predicted

F;79-113/Disulfide bonds: #status predicted

Query Match 91.7%; Score 22; DB 2; Length 196;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 158 GTTPPS 163

RESULT 8

T51838
blue copper binding protein homolog [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000
C;Accession: T51838
R;Yang, K.Y.; Kim, C.S.; Cho, B.H.
submitted to the EMBL Data Library; October 1998
A;Description: Characterization of a wound-inducible Arabidopsis gene encoding a protein
A;Reference number: Z23481
A;Accession: T51838
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-196 <YAN>
A;Cross-references: EMBL:Y18227; PIDN:CRA77089.1
C;Genetics:
A;Gene: AWI 32
C;Superfamily: plastocyanin
C;keywords: copper binding

Query Match 91.7%; Score 22; DB 2; Length 196;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 158 GTTPPS 163

RESULT 9

B82133
probable outer membrane lipoprotein Slp VC1987 [imported] - Vibrio cholerae (strain N169)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 23-Mar-2001
C;Accession: B82133
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: B82133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-202 <HEI>
A;Cross-references: GB:AE004273; GB:AE003852; NID:g9656517; PIDN:AAF95135.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1987
A;Map position: 1
C;Superfamily: rnd protein

Query Match 91.7%; Score 22; DB 2; Length 202;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 134 GTTPPS 139

RESULT 10

A83345
probable thioesterase PA2411 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: A83345
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: A83345
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-254 <STO>
A;Cross-references: GB:AE004668; GB:AE004091; NID:g9948446; PIDN:AAG05799.1; GSPDB:GN
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2411
C;Superfamily: oleoyl-[acyl-carrier-protein] hydrolase; oleoyl-[acyl-carrier-protein]

Query Match 91.7%; Score 22; DB 2; Length 254;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 108 GTAAPS 113

RESULT 11

E42725
nitrite hydratase 3'-hypothetical protein orfE - Pseudomonas chlororaphis (strain B23)
C;Species: Pseudomonas chlororaphis
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 08-Oct-1999
C;Accession: E42725
R;Nishiyama, M.; Horinouchi, S.; Kobayashi, M.; Nagasawa, T.; Yamada, H.; Beppu, T.
J. Bacteriol. 173, 2465-2472, 1991
A;Title: Cloning and characterization of genes responsible for metabolism of nitrile
A;Reference number: A42725; MUID:91193202
A;Accession: E42725
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-255 <NIS>
A;Cross-references: GB:D90216; NID:g216850; PIDN:BAA14248.1; PID:d1014953; PID:g21602

Query Match 91.7%; Score 22; DB 2; Length 255;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|||
Db 103 GTAAPS 108

RESULT 12

S47325
myob protein - zebra fish
C;Species: Brachydanio rerio (zebra fish)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Dec-2000
C;Accession: S47325
R;Weinberg, E.S.
submitted to the EMBL Data Library, September 1994
A;Description: Developmental regulation of zebrafish myo D in wild type, no tail, and
A;Reference number: S47325
A;Accession: S47325
A;Status: preliminary
A;Molecule type: mRNA

A:Residues: 1-275 <WEI>
 A:Cross-references: EMBL:Z36945; NID:G535814; PID:G535815
 C:Superfamily: human myogenin

Query Match 91.7%; Score 22; DB 2; Length 275;
 Best Local Similarity 66.7%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 253 GTTAPS 258

RESULT 13

H71320
 hypothetical protein TP0462 - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C:Accession: H71320
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khatak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
 nald, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770
 A:Accession: H71320
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-280 <COL>
 A:Cross-references: GB:AE001223; GB:AE000520; NID:G3322745; PID:G332275
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0462

Query Match 91.7%; Score 22; DB 2; Length 280;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 116 GTTAPS 121

RESULT 14

H83356
 hypothetical protein AT4g30510 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H83356
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
 g Nature 408, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488
 A:Accession: H83356
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-285 <STO>
 A:Cross-references: GB:NC_001268; NID:G7269952; PID:NCAB79769.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4g30510
 A:Map position: 4

Query Match 91.7%; Score 22; DB 2; Length 285;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 233 GTSSPS 238

RESULT 15

T17957
 hypothetical protein A454L - Chlorella virus PBCV-1
 C:Species: Chlorella virus PBCV-1
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T17957
 R:Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A:Reference number: Z18806
 A:Accession: T17957
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-289 <GRA>
 A:Cross-references: EMBL:U42580; NID:q4028896; PID:AA096822.1
 A:Experimental source: specific host Chlorella strain NC64
 C:Genetics:
 A:Note: A454L
 C:Superfamily: Chlorella virus PBCV-1 hypothetical protein A454L

Query Match 91.7%; Score 22; DB 2; Length 289;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 271 GTATPS 276

RESULT 16

T34553
 hypothetical protein DKFP434L1435.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 03-Nov-2000
 C:Accession: T34553
 R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, October 1999
 A:Reference number: Z21540
 A:Accession: T34553
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-316 <POU>
 A:Cross-references: EMBL:AL122037
 A:Experimental source: adult testis; clone DKFP434L1435
 C:Genetics:
 A:Note: DKFP434L1435.1
 C:Superfamily: valine--tRNA ligase

Query Match 91.7%; Score 22; DB 2; Length 316;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 217 GTAAPS 222

RESULT 17

S62196
 hypothetical protein 2 - Methanosarcina barkeri
 C:Species: Methanosarcina barkeri
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 22-Oct-1999
 C:Accession: S62196
 R:Vorholt, J.A.; Vaupel, M.; Thauer, R.K.
 Eur. J. Biochem. 236, 309-317, 1996
 A:Title: A polyferredoxin with eight [4Fe-4S] clusters as a subunit of molybdenum for
 A:Reference number: S62194; MUID:96184912
 A:Accession: S62196
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-319 <VOR>

A:Cross-references: EMBL:X93084; NID:g1124956; PIDN:CAA63627.1; PID:e212293; PID:g112495
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

Query Match 91.7%; Score 22; DB 2; Length 319;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 66 GTSPSPS 71

RESULT 18

JN0013

synaptic vesicle membrane protein VAR-1 - Pacific electric ray
C:Species: Torpedo californica (Pacific electric ray)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Mar-2000

C:Accession: JN0013

R:Linial, M.; Miller, K.; Scheller, R.H.

Neuron 2, 1265-1273, 1989

A:Title: VAR-1: an abundant membrane protein from torpedo cholinergic synaptic vesicles.

A:Reference number: JN0013; MUID:90166593

A:Accession: JN0013

A:Molecule type: mRNA

A:Residues: 1-379 <LIN>

C:Comment: Synaptic vesicles are responsible for regulating the storage and release of neurotransmitters.

C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

C:Keywords: glycoprotein; membrane protein; phosphoprotein

F:52-350/Domain: long-chain alcohol dehydrogenase homology <LADH>

F:67,127,147/Binding site: carboxylate (Asp) (covalent) #status predicted

F:273/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 91.7%; Score 22; DB 2; Length 379;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 194 GTSPSPS 199

RESULT 19

S12223

naringenin-chalcone synthase (EC 2.3.1.74) 1 - tomato

C:Species: Lycopersicon esculentum (tomato)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000

C:Accession: S12223

R:O'Neill, S.D.; Tong, Y.; Spoerlein, B.; Forkmann, G.; Yoder, J.I.

Mol. Gen. Genet. 224, 279-288, 1990

A:Title: Molecular genetic analysis of chalcone synthase in Lycopersicon esculentum and Lycopersicon peruvianum

A:Reference number: S12223; MUID:91117196

A:Accession: S12223

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-381 <ONE>

C:Genetics:

A:Map position: 9

C:Superfamily: chalcone synthase

C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 2; Length 381;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 23 GTSPSPS 28

RESULT 20

JN0013

naringenin-chalcone synthase (EC 2.3.1.74) 2 - potato

C:Species: Solanum tuberosum (potato)

C:Accession: JN0013

C:Date: 31-Mar-1990

C:Accession: JN0013

E87549

type IV secretion system protein B10, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: E87549

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko-

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87549

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-383 <STO>

A:Cross-references: GB:AE005673; NID:g13423963; PIDN:AAK24393.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2422

C:Superfamily: tumor-inducing plasmid pTIC58 virB10 protein

Query Match 91.7%; Score 22; DB 2; Length 383;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 74 GTSPSPS 79

RESULT 21

SYPUCJ

naringenin-chalcone synthase (EC 2.3.1.74) J - garden petunia

N:Alternate names: chalcone synthase

C:Species: Petunia x hybrida (garden petunia)

C:Date: 30-Sep-1991 #sequence_revision 17-Mar-2000 #text_change 05-May-2000

C:Accession: D72821; JS0309

R:Koes, R.B.; Spelt, C.E.; van den Elzen, P.J.M.; Mol, J.N.M.

Gene 81, 245-257, 1989

A:Title: Cloning and molecular characterization of the chalcone synthase multigene fa

A:Reference number: JS0308; MUID:90034197

A:Accession: D72821

A:Molecule type: DNA

A:Residues: 1-389 <KOE>

A:Cross-references: EMBL:X14597; NID:g20535; PIDN:CAA32737.1; PID:g20536

A:Experimental source: strain Violet 30, leaf

A:Accession: JS0309

A:Molecule type: DNA

A:Residues: 1-50, 'D', 52-74, 'V', 76-228, 'I', 230-297, 'L', 299-389 <KOE>

A:Note: the sequence is revised in GenBank entry PHCHSJ release 114, (PIDN:CAA32737.1

C:Comment: This enzyme plays a central role in the biosynthesis of all classes of fla

C:Genetics:

A:Gene: chsJ

A:Map position: V

A:Note: chsJ is expressed in various floral tissues and UV illuminated seedlings

C:Superfamily: chalcone synthase

C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 23 GTAPPS 28

RESULT 22

JCS136

naringenin-chalcone synthase (EC 2.3.1.74) 2 - potato

N:Alternate names: chalcone synthase; CHS

C:Species: Solanum tuberosum (potato)

C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 05-May-2000
 C;Accession: JC5136; PC4239
 R;Jeon, J.H.; Kim, H.S.; Choi, K.H.; Jung, Y.H.; Jung, H.; Byun, S.M.
 Biosci. Biotechnol. Biochem. 60, 1907-1910, 1996
 A;Title: Cloning and characterization of one member of the chalcone synthase gene family
 A;Reference number: JC5136; MUID:97141614
 A;Accession: JC5136
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-389 <JB01>
 A;Cross-references: GB:U47738; NID:g1470059; PIDN:AAB05239.1; PID:g1470060
 A;Accession: PC4239
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 158-165;367-373 <JB02>
 A;Comment: This enzyme is important in the biosynthesis of all classes of flavonoids in
 C;Superfamily: chalcone synthase
 C;Keywords: acyltransferase; coenzyme A

Query Match 91.7%; Score 22; DB 2; Length 389;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 23 GTATPS 28

RESULT 23

S42523
 naringenin-chalcone synthase (EC 2.3.1.74) - parsley
 N;Alternate names: chalcone synthase
 C;Species: Petroselinum crispum (parsley)
 C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 05-May-2000
 C;Accession: S42523
 R;Reimold, U.; Kroegeer, M.; Kreuzaler, F.; Hahlbrock, K.
 EMBO J. 2, 1801-1805, 1983
 A;Title: Coding and 3' non-coding nucleotide sequence of chalcone synthase mRNA and assi
 A;Reference number: S42523
 A;Accession: S42523
 A;Molecule type: mRNA
 A;Residues: 1-398 <REI>
 A;Cross-references: EMBL:V01538; NID:g20513; PIDN:CAA24779.1; PID:g20514
 A;Note: the source is designated as Petroselinum hortense
 C;Superfamily: chalcone synthase
 C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 2; Length 398;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 28 GTATPS 33

RESULT 24

S12224
 naringenin-chalcone synthase (EC 2.3.1.74) 2 - tomato
 C;Species: Lycopersicon esculentum (tomato)
 C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000
 C;Accession: S12224
 R;O'Neill, S.D.; Tong, Y.; Spoerlein, B.; Forkmann, G.; Yoder, J.I.
 Mol. Gen. Genet. 224, 279-288, 1990
 A;Title: Molecular genetic analysis of chalcone synthase in Lycopersicon esculentum and
 A;Reference number: S12223; MUID:91117196
 A;Accession: S12224
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-410 <ONE>
 C;Genetics:

A;Map position: 5
 C;Superfamily: chalcone synthase
 C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 2; Length 410;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 23 GTATPS 28

RESULT 25

F75587
 probable glycosyltransferase - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C;Accession: F75587
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A;Reference number: A75250; MUID:20036896
 A;Accession: F75587
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-418 <WHI>
 A;Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12451.1; PID:g646
 A;Experimental source: strain R1
 C;Genetics:

A;Map position: 2

C;Superfamily: glycosyltransferase

Query Match 91.7%; Score 22; DB 2; Length 418;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 365 GTATPS 370

RESULT 26

SVBJCD
 naringenin-chalcone synthase (EC 2.3.1.74) D - garden petunia
 N;Alternate names: chalcone synthase
 C;Species: Petunia x hybrida (garden petunia)
 C;Date: 30-Sep-1991 #sequence_revision 17-Mar-2000 #text_change 21-Jul-2000
 C;Accession: A72821; JS0312
 R;Koes, R.E.; Spelt, C.E.; van den Elzen, P.J.M.; Mol, J.N.M.
 Gene 81, 245-257, 1989
 A;Title: Cloning and molecular characterization of the chalcone synthase multigene fa
 A;Reference number: JS0308; MUID:90034197
 A;Accession: A72821
 A;Molecule type: DNA
 A;Residues: 1-419 <KOE>
 A;Cross-references: GB:X14593; NID:g20528; PIDN:CAA32733.1; PID:g20529
 A;Experimental source: strain Violet 30, leaf
 A;Accession: JS0312
 A;Molecule type: DNA
 A;Residues: 1-117,'D',119,'H',121-313,'DI',316-419 <KO2>
 A;Note: the sequence was revised in GenBank entry PRCHSD, release 114, (PIDN:CAA32733
 C;Comment: this enzyme plays a central role in the biosynthesis of all classes of fla
 C;Genetics:

A;Map position: V

C;Superfamily: chalcone synthase

C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 1; Length 419;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||

Db 23 GTATPS 28

RESULT 27

T36272

hypothetical protein SCE68.22 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36272

R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21576

A:Accession: T36272

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-419 <MUR>

A:Cross-references: EMBL:AL079345; PIDN:CAB45358.1; GSPDB:GN00070; SCOEDB:SCE68.22

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCE68.22

Query Match

91.7%; Score 22; DB 2; Length 419;

Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||

Db 380 GTASPS 385

RESULT 28

T41156

probable WD repeat transcription regulation protein - fission yeast (Schizosaccharomyces

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999

C:Accession: T41156

R:Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z21973

A:Accession: T41156

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-421 <HIL>

A:Cross-references: EMBL:AL031907; PIDN:CAA21427.1; GSPDB:GN00068; SPDB:SPCC18.13

A:Experimental source: strain 972h-; cosmid c18

C:Genetics:

A:Gene: SPDB:SPCC18.13

A:Map position: 3

A:Introns: 227/2

Query Match

91.7%; Score 22; DB 2; Length 421;

Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||

Db 410 GTSSPS 415

RESULT 29

T24477

hypothetical protein T04H1.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T24477

R:Harris, B.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19896

A:Accession: T24477

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-435 <WIL>

A:Cross-references: EMBL:Z78200; PIDN:CAB01584.1; GSPDB:GN00023; CESP:T04H1.7

A:Experimental source: clone T04H1

C:Genetics:

A:Gene: CESP:T04H1.7

A:Map position: 5

A:Introns: 54/3; 82/3; 193/1; 239/3; 282/3; 360/1

C:Superfamily: glucuronosyltransferase

Query Match

91.7%; Score 22; DB 2; Length 435;

Best Local Similarity 66.7%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||

Db 300 GTTTPS 305

RESULT 30

S34969

outer membrane porin OprE precursor PA0291 [imported] - Pseudomonas aeruginosa (strai

C:Species: Pseudomonas aeruginosa

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 15-Jun-2001

C:Accession: S34969; E83608

R:Yamano, Y.; Nishikawa, T.; Komatsu, Y.

Mol. Microbiol. 8, 993-1004, 1993

A:Title: Cloning and nucleotide sequence of anaerobically induced porin protein E1 (O

A:Reference number: S34969; MUID:93360827

A:Accession: S34969

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-460 <YAM>

A:Cross-references: GB:D12711; PIDN:BA02207.1; PID:g433418

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adnan, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: E83608

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-460 <STO>

A:Cross-references: GB:AE004467; GB:AE004091; MID:g9946133; PIDN:AAG03680.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: oprE; PA0291

C:Superfamily: Pseudomonas aeruginosa outer membrane porin PA0958

Query Match

91.7%; Score 22; DB 2; Length 460;

Best Local Similarity 66.7%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||

Db 54 GTASPS 59

RESULT 31

T23087

hypothetical protein H13N06.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T23087

R:Lennard, N.

Submitted to the EMBL Data Library, October 1997

A:Reference number: Z19673

A:Accession: T23087

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-460 <WTL>

A:Cross-references: EMBL:Z99942; PIDN:CAB17068.1; GSPDB:GN00028; CESP:H13N06.2

A:Experimental source: clone H13N06

C:Genetics:

A:Gene: CESP:H13N06.2

A:Map position: X

A:Introns: 17/1; 294/1; 331/3; 364/1; 414/2; 432/2

Query Match 91.7%; Score 22; DB 2; Length 460;

Best Local Similarity 66.7%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 86 GTSTPS 91

RESULT 32

T26809

hypothetical protein Y66A7A.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T26809; T27293

R:Steward, C.

Submitted to the EMBL Data Library, October 1998

A:Reference number: Z20269

A:Accession: T26809

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-507 <WIL>

A:Cross-references: EMBL:AL032627; PIDN:CAA21552.1; CESP:Y66A7A.8

A:Experimental source: clone Y41C4A

R:Matthews, L.

Submitted to the EMBL Data Library, October 1998

A:Reference number: Z20338

A:Accession: T27293

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-507 <W12>

A:Cross-references: EMBL:AL032622; PIDN:CAA21503.1; GSPDB:GN00021; CESP:Y66A7A.8

A:Experimental source: clone Y66A7A

C:Genetics:

A:Gene: CESP:Y66A7A.8

A:Map position: 3

A:Introns: 29/3; 93/1; 145/2; 314/3; 358/3

Query Match

Best Local Similarity 91.7%; Score 22; DB 2; Length 507;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 425 GTSSPS 430

RESULT 33

F75460

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: F75460

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.W.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: F75460

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-518 <WH1>

A:Cross-references: GB:AE001943; GB:AE000513; NID:g6458624; PIDN:AAF10486.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0903

A:Map position: 1

Query Match

Best Local Similarity 91.7%; Score 22; DB 2; Length 518;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 179 GTTTPS 184

RESULT 34

T40520

hypothetical protein SPBC530.04 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T40520

R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.

Submitted to the EMBL Data Library, May 1998

A:Reference number: Z21934

A:Accession: T40520

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-522 <LYN>

A:Cross-references: EMBL:AL023634; PIDN:CAA19170.1; GSPDB:GN00067; SPDB:SPBC530.04

A:Experimental source: strain 972h-; cosmid c530

C:Genetics:

A:Gene: SPDB:SPBC530.04

A:Map position: 2

A:Introns: 20/3

Query Match

Best Local Similarity 91.7%; Score 22; DB 2; Length 522;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 41 GTSAPS 46

RESULT 35

T48488

hypothetical protein T28J14.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T48488

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.;

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T48488

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-542 <BEV>

A:Cross-references: EMBL:AL163652

A:Experimental source: cultivar Columbia; BAC clone T28J14

C:Genetics:

A:Map position: 5

A:Introns: 162/3; 227/1; 251/1; 299/2; 355/3; 377/2; 397/3; 419/3; 435/3; 449/2; 487/

A>Note: T28J14.110

Query Match 91.7%; Score 22; DB 2; Length 542;
 Best Local Similarity 66.7%; Pred. No. 3.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 400 GTTSPS 405

RESULT 36
 B70583
 probable transposase for IS1535 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: B70583
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, M.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: B70583
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-550 <COL>
 A:Cross-references: GB:295210; GB:AL123456; NID:g3261757; PIDN:CAB08504.1; PID:e315222;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv0922

Query Match 91.7%; Score 22; DB 2; Length 550;
 Best Local Similarity 66.7%; Pred. No. 3.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 497 GTTAPS 502

RESULT 37
 T29005
 hypothetical protein ZK328.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29005
 R:Favell, A.
 submitted to the EMBL Data Library, March 1996
 A:Description: The sequence of C. elegans cosmid ZK328.
 A:Reference number: Z20552
 A:Accession: T29005
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-574 <FAV>
 A:Cross-references: EMBL:U50193; PIDN:AAA91246.1; CESP:ZK328.3
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:ZK328.3
 A:Introns: 63/1; 140/2; 419/1; 436/2; 488/2; 516/1; 545/2

Query Match 91.7%; Score 22; DB 2; Length 574;
 Best Local Similarity 66.7%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 284 GTSTPS 289

RESULT 38

A96783
 unknown protein F22H5.2 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96783
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, Ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talloker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: A96783
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-593 <STO>
 A:Cross-references: GB:AE005173; NID:g10092269; PIDN:AAG12682.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F22H5.2
 A:Map position: 1

Query Match 91.7%; Score 22; DB 2; Length 593;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 22 GTSSPS 27

RESULT 39
 T35746
 hypothetical protein SC7H2.15c - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T35746
 R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21588
 A:Accession: T35746
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-597 <SAU>
 A:Cross-references: EMBL:AL109732; PIDN:CAB52057.1; GSPDB:GN00070; SCOEDB:SC7H2.15c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC7H2.15c

Query Match 91.7%; Score 22; DB 2; Length 597;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 299 GTATPS 304

RESULT 40
 S20516
 dnaK-type molecular chaperone hsp70, chloroplast - chromophytic alga (Pavlova lutheri
 N:Alternate names: heat shock protein 70
 C:Species: chloroplast Pavlova lutheri
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
 C:Accession: S20516
 R:Scaramuzzi, C.D.; Stokes, H.W.; Hiller, R.G.
 Plant Mol. Biol. 18, 467-476, 1992
 A:Title: Heat shock Hsp70 protein is chloroplast-encoded in the chromophytic alga Pav

A:Reference number: S20516; MUID:92163012

A:Accession: S20516

A:Molecule type: DNA

A:Residues: 1-629 <SCA>

A:Cross-references: EMBL:X59555; NID:g20904; PIDN:CAA42154.1; PID:g20905

C:Genetics:

A:Gene: hsp70

A:Genome: chloroplast

C:Function:

A:Description: involved in protein folding and assembling/disassembling of protein complex

C:Superfamily: heat shock protein 70

C:Keywords: ATP; chloroplast; molecular chaperone

Query Match 91.7%; Score 22; DB 2; Length 629;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 33 GTTTPS 38

RESULT 41

A96781

unknown protein F9E10.5 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A96781

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: A96781

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-651 <STO>

A:Cross-references: GB:AE005173; NID:g6646755; PIDN:AAF21067.1; GSPDB:GN00141

C:Genetics:

A:Gene: F9E10.5

A:Map position: 1

Query Match 91.7%; Score 22; DB 2; Length 651;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 129 GTSSPS 134

RESULT 42

B75622

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: B75622

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A>Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: B75622

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-691 <WHI>

A:Cross-references: GB:AE001826; NID:g6460827; PIDN:AAF12628.1; PID:g6460924; TIGR:DR

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRB0037

A:Map position: megaplasmid

A:Genome: plasmid

A>Note: plasmid MP1

Query Match 91.7%; Score 22; DB 2; Length 691;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 563 GTTTPS 568

RESULT 43

T15152

hypothetical protein F35F10.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15152

R:Rohlfing, T.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid F35F10.

A:Reference number: Z18301

A:Accession: T15152

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-693 <ROH>

A:Cross-references: EMBL:AF002198; NID:g2076898; PID:g2076909; PIDN:AAB53997.1; GSPDB

A:Experimental source: strain Bristol N2; clone F35F10

C:Genetics:

A:Gene: CBSP:F35F10.10

A:Map position: 5

A:Introns: 21/2; 65/1; 83/1; 169/3; 193/3; 235/3; 259/3; 289/1; 369/3; 412/3; 495/1;

Query Match 91.7%; Score 22; DB 2; Length 693;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6

|| ||

Db 486 GTTTPS 491

RESULT 44

A29635

Transcription factor Sp1 - human (fragment)

N:Alternate names: finger protein ZNF76

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 05-Nov-1999

C:Accession: A29635; G44256

R:Kadonaga, J.T.; Carner, K.R.; Masiarz, F.R.; Tjian, R.

Cell 51, 1079-1090, 1987

A>Title: Isolation of cDNA encoding transcription factor Sp1 and functional analysis

A:Reference number: A29635; MUID:88080466

A:Accession: A29635

A:Molecule type: mRNA

A:Residues: 1-696 <KAD>

A:Cross-references: GB:J0133; NID:g339517; PIDN:AAA61154.1; PID:g339518

R:Ragoussis, J.; Senger, G.; Mockridge, I.; Sansseau, P.; Ruddy, S.; Dudley, K.; Sheer

Genomics 14, 673-679, 1992

A>Title: A testis-expressed Zn finger gene (ZNF76) in human 6p21.3 centromeric to the

A:Reference number: A44256; MUID:93052398

A:Accession: G44256

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 569-598 <RAG>
A:Experimental source: T-cell line CEM
A:Note: Sequence extracted from NCHI backbone (NCBIP:125980)
C:Genetics:
A:Gene: GDB:SP1
A:Cross-references: GDB:127453; OMIM:189906
A:Map position: 19q13.1-19q13.3
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 91.7%; Score 22; DB 2; Length 696;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 647 GTATPS 652

RESULT 45

D96527
protein F77J15.24 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96527
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dwyer, K.;
ansen, N.P.; Hughes, B.; Huizar, L.;
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ket, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: D96527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-756 <STO>
A:Cross-references: GB:AE005173; NID:g7770335; PIDN:AAF69705.1; GSPDB:GN00141
C:Genetics:
A:Gene: F27J15.24
A:Map position: 1

Query Match 91.7%; Score 22; DB 2; Length 756;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 118 GTSAPS 123

RESULT 46

JS0747
regulatory protein Spl - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C:Accession: JS0747; S25287
R:Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami,
submitted to JIPID, September 1992
A:Reference number: JS0747
A:Accession: JS0747
A:Molecule type: mRNA
A:Residues: 1-788 <IMA>
A:Cross-references: DDBJ:D12768; NID:g220911; PIDN:BAA02235.1; PID:d1002730; PID:g220912
R:Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami,
EMBO J. 11, 3663-3671, 1992
A:Title: Two regulatory proteins that bind to the basic transcription element (BTE), a G
A:Reference number: S25287; MUID:93010958
A:Accession: S25287

A:Molecule type: mRNA
A:Residues: 1-122,'L',124-311,'A',313-788 <IM2>
C:Keywords: DNA binding; transcription regulation

Query Match 91.7%; Score 22; DB 2; Length 788;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 739 GTATPS 744

RESULT 47

T28055
hypothetical protein ZK858.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28055
R:White, S.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z20462
A:Accession: T28055
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-848 <WIL>
A:Cross-references: EMBL:Z79759; PIDN:CAB02138.1; GSPDB:GN00019; CESP:ZK858.1
A:Experimental source: clone ZK858
C:Genetics:
A:Gene: CESP:ZK858.1
A:Map position: 1
A:Introns: 26/3; 82/3; 130/3; 195/3; 241/3; 613/1; 762/1

Query Match 91.7%; Score 22; DB 2; Length 848;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 630 GTAAPS 635

RESULT 48

T20429
hypothetical protein E03A3.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T20429
R:Gardner, A.
submitted to the EMBL Data Library, October 1994
A:Reference number: Z19274
A:Accession: T20429
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-875 <WIL>
A:Cross-references: EMBL:Z38112; PIDN:CAA86231.1; GSPDB:GN00021; CESP:E03A3.6
A:Experimental source: clone E03A3
C:Genetics:
A:Gene: CESP:E03A3.6
A:Map position: 3
A:Introns: 58/2; 161/3; 200/1; 257/3; 436/3; 629/3; 735/1; 762/3
C:Superfamily: Caenorhabditis elegans hypothetical protein E03A3.6

Query Match 91.7%; Score 22; DB 2; Length 875;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 827 GTSAPS 832

RESULT 49
A47521
capsid protein - giardiavirus GLV
C:Species: giardiavirus, GLV
C:Date: 21-Jan-1994 #sequence_revision 13-Feb-1998 #text_change 16-Jul-1999
C:Accession: A47521
R:Wang, A.L.; Yang, H.M.; Shen, K.A.; Wang, C.C.
Proc. Natl. Acad. Sci. U.S.A. 90, 8595-8599, 1993
A:Title: Giardiavirus double-stranded RNA genome encodes a capsid polypeptide and a gag
A:Reference number: A47521; MUID:93391401
A:Contents: host Giardia lamblia
A:Accession: A47521
A:Molecule type: genomic RNA
A:Residues: 1-886 <WAN>
A:Cross-references: GB:L13218; NID:gi1352866; PIDN:AAB01578.1; PID:gi1352867
A:Note: sequence modified after extraction from NCBI backbone (NCBIN:137593, NCBI:137593)
C:Superfamily: giardiavirus capsid protein
C:Keywords: capsid protein

Query Match 91.7%; Score 22; DB 1; Length 886;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
Db 796 GTAAPS 801

RESULT 50
T40290
hypothetical protein SPBC354.10 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40290
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21918
A:Accession: T40290
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-963 <WOO>
A:Cross-references: EMBL:AL022071; PIDN:CAA17810.1; GSPDB:GN00067; SPDB:SPBC354.10
A:Experimental source: strain 972h-; cosmid c354
C:Genetics:
A:Gene: SPDB:SPBC354.10
A:Map position: 2

Query Match 91.7%; Score 22; DB 2; Length 963;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
Db 730 GTAAPS 735

Search completed: September 24, 2002, 11:24:19
Job time: 4078 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein. - protein search, using sw model

Run on: September 24, 2002, 11:27:06 : Search time 39.04 Seconds
(without alignments)
5.951 Million cell updates/sec

Title: BASK-853-CLAIM5
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38715550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	196	1	Q07488 arabidopsis
2	22	91.7	255	1	Q03003 pseudomonas
3	22	91.7	275	1	Q90477 brachydanio
4	22	91.7	312	1	Q92269 mus musculus
5	22	91.7	379	1	P19333 torpeda cal
6	22	91.7	389	1	P23418 lycopersico
7	22	91.7	389	1	P23419 lycopersico
8	22	91.7	389	1	Q43188 solanum tub
9	22	91.7	389	1	Q41436 solanum tub
10	22	91.7	389	1	Q43163 solanum tub
11	22	91.7	389	1	P22928 petunia hyb
12	22	91.7	389	1	Q92784 catharanthu
13	22	91.7	389	1	P49440 phaseolus v
14	22	91.7	392	1	O83475 treponema p
15	22	91.7	395	1	P51075 betula verr
16	22	91.7	397	1	Q92840 daucus caro
17	22	91.7	398	1	P48385 callistephu
18	22	91.7	398	1	P16107 petroselinu
19	22	91.7	419	1	P22925 petunia hyb
20	22	91.7	435	1	Q22180 caenorhabdi
21	22	91.7	629	1	P30722 pavlova lut
22	22	91.7	785	1	P08047 homo sapien
23	22	91.7	788	1	Q01714 rattus norv
24	22	91.7	1050	1	O75385 homo sapien
25	22	91.7	1262	1	P27671 mus musculu
26	22	91.7	1377	1	P19538 drosophila
27	22	91.7	2717	1	P15822 homo sapien
28	22	91.7	2769	1	P01267 bos taurus
29	21	87.5	45	1	P05821 escherichia
30	21	87.5	45	1	LYS3_ECOLI
31	21	87.5	47	1	LYS3_SHISO
32	21	87.5	47	1	LYS2_ECOLI
33	21	87.5	47	1	LYS5_ECOLI
34	21	87.5	47	1	LYS6_ECOLI

34	21	87.5	47	1	LYS7_ECOLI	Q03709 escherichia
35	21	87.5	47	1	LYS8_ECOLI	P10099 escherichia
36	21	87.5	47	1	LYS9_ECOLI	P15176 escherichia
37	21	87.5	49	1	LYS0_ECOLI	P02987 escherichia
38	21	87.5	52	1	LYS4_ECOLI	P09181 escherichia
39	21	87.5	80	1	PAHQ_CHICK	P01306 gallus gall
40	21	87.5	94	1	YQGV_BACSU	P54499 bacillus su
41	21	87.5	118	1	TRHA_ECOLI	Q00190 escherichia
42	21	87.5	122	1	RK14_PORPU	P51304 porphyra pu
43	21	87.5	170	1	FMF7_ECOLI	P25394 escherichia
44	21	87.5	187	1	OM24_ARATH	P82805 arabidopsis
45	21	87.5	193	1	Y319_MYCPN	P75329 mycoplasma
46	21	87.5	203	1	COX3_MYCTU	Q10385 mycobacteri
47	21	87.5	215	1	NODB_RHISN	P50355 rhizobium s
48	21	87.5	225	1	XYN2_ASPNG	P55330 aspergillus
49	21	87.5	242	1	PDXJ_NEIMA	Q9rqv9 neisseria m
50	21	87.5	242	1	PDXJ_NEIMB	Q9k0v9 neisseria m

ALIGNMENTS

RESULT 1

BCB_ARATH

ID BCB_ARATH STANDARD; PRT; 196 AA.
AC Q07488;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Blue copper protein precursor.

GN BCB.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLOMBIA, AND CV. K85;

RX MEDLINE=94124044; PubMed=8294044;

RA van Gysel A.; van Montagu M.; Inze D.;

RT "A negatively light-regulated gene from Arabidopsis thaliana encodes

a protein showing high similarity to blue copper-binding proteins.";

RL Gene 136:79-85(1993).

CC -!- FUNCTION: PROBABLY ACTS AS AN ELECTRON CARRIER.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -!- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND IN 35 DAY OLD

PLANTLETS WHEN THE ROSETTE IS MATURE, CONSISTING OF 8-10 FULLY

EXPANDED LEAVES, AND AS THE FLORAL STEM STARTS TO FORM. THIS LEVEL

REMAINS CONSTANT DURING THE FURTHER LIFE SPAN OF THE PLANT.

CC -!- INDUCTION: BY DARK ADAPTATION. THIS GIVES A 20-FOLD INCREASE IN

EXPRESSION.

CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z15058; CAA78771.1; -

CC PIR; S25555; S25555.

DR HSSP; P29602; 1JER.

DR InterPro; IPR000923; Copper_blue1.

DR InterPro; IPR003245; Cu_bind_like.

DR Pfam; PF02298; Cu_bind_like; 1.

DR ProDom; PD003122; Cu_bind_like; 1.

DR PROSITE; PS00196; COPPER_BLUE; 1.

KW Electron transport; Copper; Transmembrane; Signal; Metal-binding;

KW Glycoprotein. 1 22 POTENTIAL.

FT SIGNAL

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98409499; PubMed=9736772;
RA Laporte J., Blondeau F., Buj-Bello A., Tentler D., Kretz C., Dahl N.,
Mandel J.-L.;
RT "Characterization of the myotubularin dual specificity phosphatase
gene family from yeast to human."
RL Hum. Mol. Genet. 7:1703-1712(1998).
CC -1- FUNCTION: NOT KNOWN, COULD BE A TYROSINE-PHOSPHATASE.
CC -1- SIMILARITY: BELONGS TO THE MYOTUBULARIN-LIKE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF073882; AAC80004.1; -
DR MGD; MGI:1891693; Mtmr7.
DR InterPro: IPR000340; DS_phosphatase.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
FT NON_TER 1 1
FT ACT_SITE 48 48 BY SIMILARITY.
SQ SEQUENCE 312 AA; 36018 MW; 4BB40E0D15BD880C CRC64;

Query Match 91.7%; Score 22; DB 1; Length 312;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
| | | |
Db 294 GTASPS 299

RESULT 5
VAT1_TORCA STANDARD; PRT; 379 AA.
AC P19333;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Synaptic vesicle membrane protein VAT-1.
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hyncoisalea; Pristiorajea; Batoidae;
OC Torpediniformes; Torpedinoidei; Torpedinidae; Torpedo.
OX NCBI_TaxID=7787;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=electric lobe;
RX MEDLINE=90166593; PubMed=2483112;
RA Linial M., Miller K., Scheller R.H.;
RT "VAT-1: an abundant membrane protein from Torpedo cholinergic
synaptic vesicles."
RL Neuron 2:1265-1273(1989).
CC -1- FUNCTION: MAY PLAY A CENTRAL ROLE IN THE FUNCTIONS MEDIATED BY
CC SPECIFIC CLASSES OF SYNAPTIC VESICLES.
CC -1- SUBCELLULAR LOCATION: Membrane.
CC -1- TISSUE SPECIFICITY: CHOLINERGIC SYNAPTIC VESICLES.
CC -1- MISCELLANEOUS: SYNTHESIZED IN THE NEURONAL CELL BODIES AND
CC TRANSPORTED TO THE TERMINALS. EACH VESICLE CONTAINS APPROXIMATELY
CC 28 MOLECULES OF VAT-1.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE

CC FAMILY. QUINONE OXIDOREDUCTASE SUBFAMILY. STRONG, TO MAMMALIAN
CC VAT-1 HOMOLOGS.
DR PIR; JN0013; JN0013.
DR InterPro: IPR002085; Adh_zn_family.
DR Pfam; PF002364; QOR_zeta_crystal.
DR PROSITE; PS01162; adh_zinc; 1.
DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
KW Oxidoreductase; Zinc; Synapse; Membrane; Phosphorylation.
FT MOD_RES 273 273 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 379 AA; 41593 MW; FA4ADA17E657F09C CRC64;

Query Match 91.7%; Score 22; DB 1; Length 379;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
| | | |
Db 194 GTASPS 199

RESULT 6
CHS1_LYCES STANDARD; PRT; 389 AA.
ID CHS1_LYCES
AC P23418;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).
GN CHS1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon, Hypocotyl, and Leaf;
RX MEDLINE=91117196; PubMed=1980524;
RA O'Neill S.D., Tong Y., Spoerlein B., Forkmann G., Yoder J.I.;
RT "Molecular genetic analysis of chalcone synthase in Lycopersicon
esculentum and an anthocyanin-deficient mutant."
RL Mol. Gen. Genet. 224:279-288(1990).
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 Co(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X55194; CAA38980.1; -
DR PIR; S12223; S12223.
DR InterPro: IPR001099; Chal_stil_synt.
DR Pfam; PF00195; Chal_stil_synt; 1.
DR Pfam; PF02797; Chal_stil_syntC; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;
KW Multigene family.
FT ACT_SITE 164 164
SQ SEQUENCE 389 AA; 42552 MW; 553DC69E5EA96A8B CRC64;

```

Query Match          91.7%; Score 22; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 23 GTSTPS 28

RESULT 7
CHS2_LYCES
ID CHS2_LYCES STANDARD; PRT; 389 AA.
AC P23419;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
GN CHS2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
NCBI_TaxID=4081;
[1]
RN
RP SEQUENCE FROM N.A.
RT TISSUE=Cotyledon, Hypocotyl, and Leaf;
RX MEDLINE=91117196; PubMed=1980524;
RA O'Neill S.D., Tong Y., Spoerlein B., Forkmann G., Voder J.I.;
RT "Molecular genetic analysis of chalcone synthase in Lycopersicon
RT esculentum and an anthocyanin-deficient mutant.";
RL Mol. Gen. Genet. 224:279-288(1990).
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X55195; CAA38981.1; -.
CC PIR; S12224; S12224.
CC InterPro; IPR001099; Chal_stil_synt.
CC Pfam; PF00195; Chal_stil_synt; 1.
CC Pfam; PF02797; Chal_stil_syntC; 1.
CC PROSITE; PS00441; CHALCONE_SYNTH; 1.
CC Flavonoid biosynthesis; Transferase; Acyltransferase;
CC Multigene family.
CC ACT_SITE 164 164 BY SIMILARITY.
CC SEQUENCE 389 AA; 42730 MW; F92E46BB3B5FC32F CRC64;

Query Match          91.7%; Score 22; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 23 GTATPS 28

RESULT 8

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-CV. RED PONTIAC;
RA Jeon J.-H., Kim H.-S., Choi K.-H., Joung Y.-H., Joung H., Byun S.-M.;
RT "Characterization of two members of the chalcone synthase gene family
RT from Solanum tuberosum L.";
RL (In) Plant Gene Register PGR96-027.
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U47739; AAB67734.1; -
CC Mendel; 10624; SOLtu;1193;1.
CC InterPro: IPR001099; Chal_stil_synth.
CC Pfam: PF00195; Chal_stil_synth; 1.
CC Pfam: PF02797; Chal_stil_synthC; 1.
CC PROSITE: PS00441; CHALCONE_SYNTH; 1.
CC Flavonoid biosynthesis; Transferase; Acyltransferase;
CC Multigene family.
CC ACT_SITE 164 BY SIMILARITY.
CC SEQUENCE 389 AA; 42562 MW; B181D9C6B9170A34 CRC64;
CC -----
CC Query Match 91.7%; Score 22; DB 1; Length 389;
CC Best Local Similarity 66.7%; Pred. No. 92;
CC Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 1 gtxxps 6
CC | | | |
CC | | | |
CC Db 23 GTSTPS 28
CC -----
RESULT 10
CHSB_SOLTU
ID CHSB_SOLTU STANDARD; PRT; 389 AA.
AC Q43163
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 1B (EC 2.3.1.74) (Naringenin-chalcone synthase 1B).
GN CHSLB.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. RED PONTIAC;
RA Jeon J.-H., Kim H.-S., Choi K.-H., Joung Y.-H., Joung H., Byun S.-M.;
RT "Characterization of two members of the chalcone synthase gene family
RT from Solanum tuberosum L.";
RL (In) Plant Gene Register PGR96-027.
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U47739; AAB67734.1; -
CC Mendel; 10624; SOLtu;1193;1.
CC InterPro: IPR001099; Chal_stil_synth.
CC Pfam: PF00195; Chal_stil_synth; 1.
CC Pfam: PF02797; Chal_stil_synthC; 1.
CC PROSITE: PS00441; CHALCONE_SYNTH; 1.
CC Flavonoid biosynthesis; Transferase; Acyltransferase;
CC Multigene family.
CC ACT_SITE 164 BY SIMILARITY.
CC SEQUENCE 389 AA; 42562 MW; B181D9C6B9170A34 CRC64;
CC -----
CC Query Match 91.7%; Score 22; DB 1; Length 389;
CC Best Local Similarity 66.7%; Pred. No. 92;
CC Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 1 gtxxps 6
CC | | | |
CC | | | |
CC Db 23 GTSTPS 28
CC -----
RESULT 11
CHSJ_PETHY
ID CHSJ_PETHY STANDARD; PRT; 389 AA.
AC P22928
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase J (EC 2.3.1.74) (Naringenin-chalcone synthase J).
GN CHSJ.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. VIOLET 30; TISSUE-Leaf;
RX MEDLINE=90034197; PubMed=2806915;
RA Koes R.E., Spelt C.E., van den Elzen P.J.M., Mol J.N.M.;
RT "Cloning and molecular characterization of the chalcone synthase
RT multigene family of Petunia hybrida.";
RL Gene 81:245-257(1989).
RN [2]
RP SEQUENCE OF 71-389 FROM N.A.
RC STRAIN-CV. WHITE 137; TISSUE-Anther;
RA van Tunen A.J.;
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U47740; AAB67735.1; -
CC Mendel; 10625; SOLtu;1193;2.
CC InterPro: IPR001099; Chal_stil_synth.
CC Pfam: PF00195; Chal_stil_synth; 1.
CC Pfam: PF02797; Chal_stil_synthC; 1.
CC PROSITE: PS00441; CHALCONE_SYNTH; 1.
CC Flavonoid biosynthesis; Transferase; Acyltransferase;
CC Multigene family.
CC ACT_SITE 164 BY SIMILARITY.
CC SEQUENCE 389 AA; 42548 MW; E7334A1A34BD1CC CRC64;
CC -----
CC Query Match 91.7%; Score 22; DB 1; Length 389;
CC Best Local Similarity 66.7%; Pred. No. 92;
CC Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 1 gtxxps 6
CC | | | |
CC | | | |
CC Db 23 GTSTPS 28
CC -----
RESULT 11
CHSJ_PETHY
ID CHSJ_PETHY STANDARD; PRT; 389 AA.
AC P22928
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase J (EC 2.3.1.74) (Naringenin-chalcone synthase J).
GN CHSJ.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. VIOLET 30; TISSUE-Leaf;
RX MEDLINE=90034197; PubMed=2806915;
RA Koes R.E., Spelt C.E., van den Elzen P.J.M., Mol J.N.M.;
RT "Cloning and molecular characterization of the chalcone synthase
RT multigene family of Petunia hybrida.";
RL Gene 81:245-257(1989).
RN [2]
RP SEQUENCE OF 71-389 FROM N.A.
RC STRAIN-CV. WHITE 137; TISSUE-Anther;
RA van Tunen A.J.;
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation--
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X14597; CAA32737.1; -;
 DR EMBL; X14599; CAA32739.1; -;
 DR PIR; J03039; SYPICJ;
 DR InterPro; IPR001099; Chal_stil_synt.
 DR Pfam; PF00195; Chal_stil_synt; 1.
 DR Pfam; PF02797; Chal_stil_syntC; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase;
 KW Multigene family.
 FT ACT_SITE 164 164 BY SIMILARITY.
 FT CONFLICT 75 75 E -> V (IN REF. 2).
 SQ SEQUENCE 389 AA; 42558 MW; F2B3CDD82E6FDE7D CRC64;

Query Match 91.7%; Score 22; DB 1; Length 389;
 Best Local Similarity 66.7%; Pred. No. 92;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 23 GTATPS 28

RESULT 12

CHSY_CATRO ID CHSY_CATRO STANDARD; PRT; 389 AA.
 AC Q9ZRS4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Chalcone synthase (EC 2.3.1.74) (Naringenin-chalcone synthase).
 GN CHS.
 OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
 OC Vinceae; Catharanthus.
 OX NCBI_TaxID=4058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kaltenbach M., Schroeder G., Schmelzer E., Lutz V., Schroeder J.;
 RT "Flavonoid hydroxylase from Catharanthus roseus: cDNA, heterologous
 RT expression, enzyme properties, and cell-type specific expression in
 RT plants.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHIDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ131813; CAA10511.1; -;
 DR Mendel; 36879; Catro; I193; 36879.

DR InterPro; IPR001099; Chal_stil_synt.
 DR Pfam; PF00195; Chal_stil_synt; 1.
 DR Pfam; PF02797; Chal_stil_syntC; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase.
 KW Multigene family.
 FT ACT_SITE 164 164 BY SIMILARITY.
 SQ SEQUENCE 389 AA; 42492 MW; B350742DE4A19186 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 389;
 Best Local Similarity 66.7%; Pred. No. 92;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 23 GTSTPS 28

RESULT 13

CHSY_PHAVU ID CHSY_PHAVU STANDARD; PRT; 389 AA.
 AC P49440;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chalcone synthase 17 (EC 2.3.1.74) (Naringenin-chalcone synthase 17).
 GN CHS17.
 OS Phaseolus vulgaris (Kidney bean) (French bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
 OX NCBI_TaxID=3885;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CANADIAN WONDER;
 RX MEDLINE=88142539; PubMed=3481420;
 RA Ryder T.B., Hedrick S.A., Bell J.N., Liang X., Clouse S.D.,
 RA Lamb C.J.;
 RT "Organization and differential activation of a gene family encoding
 RT the plant defense enzyme chalcone synthase in Phaseolus vulgaris.";
 RL Mol. Gen. Genet. 210:219-233(1987).
 CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHIDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -!- INDUCTION: BY IRRADIATION, ELICITOR, INFECTION OR WOUNDING.
 CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X06411; CAA29700.1; -;
 DR InterPro; IPR001099; Chal_stil_synt.
 DR Pfam; PF00195; Chal_stil_synt; 1.
 DR Pfam; PF02797; Chal_stil_syntC; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase;
 KW Multigene family.
 FT ACT_SITE 164 164 BY SIMILARITY.
 SQ SEQUENCE 389 AA; 42654 MW; 25F13C3CD654F47 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 389;

```
Best Local Similarity 66.7%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 23 GTATPS 28

RESULT 14
Y462_TREPA
ID Y462_TREPA STANDARD; PRT; 392 AA.
AC O83475; O83476;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein TP0462/TP0463 precursor.
GN TP0462/TP0463.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. REF.1 SEQUENCE DIFFERS
CC FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 279.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001223; AAC65455.1; ALT_FRAME.
CC EMBL; AE001223; AAC65456.1; ALT_FRAME.
CC TIGR; TP0462; -.
CC TIGR; TP0463; -.
CC PROSITE; PS000013; PROKAR_LIPOPROTEIN; 1.
CC Hypothetical protein; Lipoprotein; Membrane; Signal;
CC Complete proteome.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 392 HYPOTHETICAL LIPOPROTEIN TP0462/TP0463.
CC LIPID 20 120 N-ACYL DIGLYCERIDE (POTENTIAL).
CC DOMAIN 148 163 GLY/SER-RICH.
CC SEQUENCE 392 AA; 40261 MW; F2689260172607AE CRC64;

Query Match 91.7%; Score 22; DB 1; Length 392;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 116 GTTTPS 121

RESULT 15
CHSY_BETVE
ID Y462_TREPA STANDARD; PRT; 392 AA.
AC O83475; O83476;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein TP0462/TP0463 precursor.
GN TP0462/TP0463.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. REF.1 SEQUENCE DIFFERS
CC FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 279.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001223; AAC65455.1; ALT_FRAME.
CC EMBL; AE001223; AAC65456.1; ALT_FRAME.
CC TIGR; TP0462; -.
CC TIGR; TP0463; -.
CC PROSITE; PS000013; PROKAR_LIPOPROTEIN; 1.
CC Hypothetical protein; Lipoprotein; Membrane; Signal;
CC Complete proteome.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 392 HYPOTHETICAL LIPOPROTEIN TP0462/TP0463.
CC LIPID 20 120 N-ACYL DIGLYCERIDE (POTENTIAL).
CC DOMAIN 148 163 GLY/SER-RICH.
CC SEQUENCE 392 AA; 40261 MW; F2689260172607AE CRC64;

Query Match 91.7%; Score 22; DB 1; Length 392;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 116 GTTTPS 121

RESULT 15
CHSY_BETVE
ID Y462_TREPA STANDARD; PRT; 392 AA.
AC O83475; O83476;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein TP0462/TP0463 precursor.
GN TP0462/TP0463.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. REF.1 SEQUENCE DIFFERS
CC FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 279.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001223; AAC65455.1; ALT_FRAME.
CC EMBL; AE001223; AAC65456.1; ALT_FRAME.
CC TIGR; TP0462; -.
CC TIGR; TP0463; -.
CC PROSITE; PS000013; PROKAR_LIPOPROTEIN; 1.
CC Hypothetical protein; Lipoprotein; Membrane; Signal;
CC Complete proteome.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 392 HYPOTHETICAL LIPOPROTEIN TP0462/TP0463.
CC LIPID 20 120 N-ACYL DIGLYCERIDE (POTENTIAL).
CC DOMAIN 148 163 GLY/SER-RICH.
CC SEQUENCE 392 AA; 40261 MW; F2689260172607AE CRC64;

Query Match 91.7%; Score 22; DB 1; Length 395;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 23 GTATPS 28

RESULT 16
CHS2_DAUCA
ID CHS2_DAUCA STANDARD; PRT; 397 AA.
AC Q9ZS40;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
GN Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX PubMed=119157 FROM N.A.
RA Pellinen R., Korhonen M., Overmyer K., Lapinjoki S., Kangasjaervi J.,
RA "Induction of different defence responses in birch (Betula pendula
RA Roth) upon ozone and UV-B stresses."
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 119-157 FROM N.A.
RC TISSUE=Leaf;
RX PubMed=119157 FROM N.A.
RA Pellinen R., Pellinen R., Roy S., Julkunen-Tiitto R., Eloranta T.,
RA Kangasjaervi J.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y11022; CAA71904.1; -.
CC EMBL; X77513; CAA54649.1; -.
CC InterPro; IPR001099; Chal_stil_synt.
CC Pfam; PF00195; Chal_stil_synt; 1.
CC Pfam; PF02797; Chal_stil_syntC; 1.
CC PROSITE; PS00441; CHALCONE_SYNTH; 1.
CC Flavanoid biosynthesis; Transferase; Acyltransferase.
CC ACT_SITE 164 164 BY SIMILARITY.
CC SEQUENCE 395 AA; 43015 MW; 18D25119A8891A17 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 395;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 23 GTATPS 28

RESULT 16
CHS2_DAUCA
ID CHS2_DAUCA STANDARD; PRT; 397 AA.
AC Q9ZS40;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
GN Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX PubMed=119157 FROM N.A.
RA Pellinen R., Korhonen M., Overmyer K., Lapinjoki S., Kangasjaervi J.,
RA "Induction of different defence responses in birch (Betula pendula
RA Roth) upon ozone and UV-B stresses."
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 119-157 FROM N.A.
RC TISSUE=Leaf;
RX PubMed=119157 FROM N.A.
RA Pellinen R., Pellinen R., Roy S., Julkunen-Tiitto R., Eloranta T.,
RA Kangasjaervi J.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y11022; CAA71904.1; -.
CC EMBL; X77513; CAA54649.1; -.
CC InterPro; IPR001099; Chal_stil_synt.
CC Pfam; PF00195; Chal_stil_synt; 1.
CC Pfam; PF02797; Chal_stil_syntC; 1.
CC PROSITE; PS00441; CHALCONE_SYNTH; 1.
CC Flavanoid biosynthesis; Transferase; Acyltransferase.
CC ACT_SITE 164 164 BY SIMILARITY.
CC SEQUENCE 395 AA; 43015 MW; 18D25119A8891A17 CRC64;
```

GN CHS2.
 OS Daucus carota (Carrot).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 CC NCBI_TaxID=4039;
 CC [1]
 RN SEQUENCE FROM N.A.
 RP Hirner A.A., Seitz H.U.;
 RA "Cloning of two different chalcone synthase isoforms from Daucus
 RT carota L. and their differential expression in organs from european
 RT wild carrot and in UV-A irradiated cell cultures.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ006780; CAA07245.1; -
 CC Mende; 36898; Daucus; 1193; 36898.
 CC InterPro; IPR001099; Chal_stil_synt.
 CC Pfam; PF00195; Chal_stil_synt; 1.
 CC Pfam; PF02797; Chal_stil_syntC; 1.
 CC PROSITE; PS00441; CHALCONE_SYNTH; 1.
 CC Flavonoid biosynthesis; Transferase; Acyltransferase;
 CC Multigene family.
 CC ACT_SITE 168 168 BY SIMILARITY.
 CC SEQUENCE 397 AA; 43559 MW; 9386F44B9132EDEE CRC64;
 DR
 DR
 DR Mendel; 36898; Daucus; 1193; 36898.
 DR InterPro; IPR001099; Chal_stil_synt.
 DR Pfam; PF00195; Chal_stil_synt; 1.
 DR Pfam; PF02797; Chal_stil_syntC; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 DR Flavonoid biosynthesis; Transferase; Acyltransferase;
 DR Multigene family.
 DR ACT_SITE 168 168 BY SIMILARITY.
 DR SEQUENCE 397 AA; 43559 MW; 9386F44B9132EDEE CRC64;
 SQ
 Query Match 91.7%; Score 22; DB 1; Length 397;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gtxxps 6
 || ||
 Db 27 GTATPS 32
 RESULT 17
 CHSY_CALCH STANDARD; PRT; 398 AA.
 ID CHSY_CALCH
 AC P48365;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chalcone synthase (EC 2.3.1.74) (Naringenin-chalcone synthase).
 GN CHS.
 OS Callistephus chinensis (China aster).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 CC Asteraceae; Callistephus.
 CC NCBI_TaxID=13379;
 CC [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-L 01; TISSUE=Petal;
 RA Henkel J., Wassenegeger M., Sommer H., Forkmann G.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z67988; CAA91930.1; -
 CC InterPro; IPR001099; Chal_stil_synt.
 CC Pfam; PF00195; Chal_stil_synt; 1.
 CC Pfam; PF02797; Chal_stil_syntC; 1.
 CC PROSITE; PS00441; CHALCONE_SYNTH; 1.
 CC Flavonoid biosynthesis; Transferase; Acyltransferase.
 CC ACT_SITE 167 167 BY SIMILARITY.
 CC SEQUENCE 398 AA; 43541 MW; 8B7D4E1B4FAFFC3 CRC64;
 SQ
 Query Match 91.7%; Score 22; DB 1; Length 398;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gtxxps 6
 || ||
 Db 26 GTATPS 31
 RESULT 18
 CHSY_PETCR STANDARD; PRT; 398 AA.
 ID CHSY_PETCR
 AC P16107;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chalcone synthase (EC 2.3.1.74) (Naringenin-chalcone synthase).
 GN CHS.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
 CC NCBI_TaxID=4043;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Reinhold U., Kroege M., Kreuzaler F., Hahlbrock K.;
 RT "Coding and 3' non-coding nucleotide sequence of chalcone synthase
 RT mRNA and assignment of amino acid sequence of the enzyme.";
 RL EMBO J. 2:1801-1805(1983).
 CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; V01538; CAA24779.1; -
 DR PIR; S42523; S42523.
 DR InterPro; IPR001099; Chal_still_synt.
 DR Pfam; PF00195; Chal_still_synt; 1.
 DR Pfam; PF02797; Chal_still_synt; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase.
 FT ACT_SITE 169
 FT ACT_SITE 169 BY SIMILARITY.
 SQ SEQUENCE 398 AA; 43735 MW; E8B19149AD3DABLE CRC64;

Query Match 91.7%; Score 22; DB 1; Length 398;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 28 GTATPS 33

RESULT 19

CHSD_PETHY STANDARD; PRT; 419 AA.
 AC P22925;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chalcone synthase D (EC 2.3.1.74) (Naringenin-chalcone synthase D).
 GN CHSD.
 OS Petunia hybrida (Petunia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
 OX NCBI_TaxID=4102;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CV. VIOLET 30; TISSUE=Leaf;
 RX MEDLINE=90034197; PubMed=2806915;
 RA Koes R.E., Spelt C.E., van den Elzen P.J.M., Mol J.N.M.;
 RT "Cloning and molecular characterization of the chalcone synthase
 multigene family of Petunia hybrida."
 RL Gene 81:245-257(1989).
 CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2'',4'',6''-
 TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 NARINGENIN
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 naringenin-chalcone + 3 Co(2).
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 OF WHICH ARE BRIGHTLY COLORED.
 CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X14593; CAA32733.1; -
 DR PIR; JS0312; SYPJCD.
 DR InterPro; IPR001099; Chal_still_synt.
 DR Pfam; PF00195; Chal_still_synt; 1.
 DR Pfam; PF02797; Chal_still_synt; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase;
 KW Multigene family.

FT ACT_SITE 164 164 BY SIMILARITY.
 SQ SEQUENCE 419 AA; 45979 MW; 55827338DC812ACD CRC64;

Query Match 91.7%; Score 22; DB 1; Length 419;
 Best Local Similarity 66.7%; Pred. No. 99;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 23 GTATPS 28

RESULT 20

UGTB_CAEEL STANDARD; PRT; 435 AA.
 ID UGTB_CAEEL
 AC Q22180;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative UDP-glucuronosyltransferase UGT11 (EC 2.4.1.17) (UDPGT).
 GN UGT11 OR T04H1.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Harris B.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: UDP-glucuronate + acceptor = UDP + acceptor
 beta-D-glucuronoside.
 CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; Z78200; CAB01584.1; -
 DR WormPep; T04H1.7; CEI3179.
 DR InterPro; IPR002213; UDPGT.
 DR Pfam; PF00201; UDPGT; 1.
 DR PROSITE; PS00375; UDPGT; 1.
 KW Hypothetical protein; Transferase; Glycosyltransferase;
 KW Multigene family.
 SQ SEQUENCE 435 AA; 50020 MW; AF04A0DE7208CBF3 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 435;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 300 GTTTPS 305

RESULT 21

DNAK_PAVLU STANDARD; PRT; 629 AA.
 ID DNAK_PAVLU
 AC P30722;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
 DE protein) (HSP70).
 DE DNAK OR HSP70.
 OS Pavlova lutherli (Monochrysis lutherli).

OG Chloroplast.
 CC Eukaryota; Haptophyceae; Pavlova; Pavlova.
 OX NCBI_TaxID-2832;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92163012; PubMed-1536924;
 RA Scaramuzzi C.D., Stokes H.W., Hiller R.G.;
 RT "Heat shock Hsp70 protein is chloroplast-encoded in the chromophytic
 RL alga Pavlova lutheri";
 CC Plant Mol. Biol. 18:467-476(1992).
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X59555; CAA42154.1; -;
 DR HSP; S20516; S20516.
 DR HSP; P04475; IDG4.
 DR InterPro; IPR001023; HSP70.
 DR Pfam; PF00012; HSP70; 1.
 DR PRINTS; PR00301; HEATSHOCK70.
 DR PROSITE; PS00297; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KW Chaperone; ATP-binding; Heat shock; Chloroplast.
 SQ SEQUENCE 629 AA; 68792 MW; C388D0C369579D66 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 629;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 DB 33 GTTTPS 38

RESULT 22
 SP1_HUMAN
 ID SP1_HUMAN STANDARD; PRT; 785 AA.
 AC P08047; Q9NVE7; Q9H3Q5;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor Sp1.
 GN SP1 OR TSFPL.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 4-785 FROM N.A.
 RC TISSUE=Cervical carcinoma;
 RA Haggart M.H., Ladurner A.G.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-558 FROM N.A.
 RX MEDLINE-20545561; PubMed-10973950;
 RA Takahara T., Kanazu S., Yanagisawa S., Akanuma H.;
 RT "Heterogeneous Sp1 mRNAs in human HepG2 cells include a product of
 RL homotypic trans-splicing";
 RN J. Biol. Chem. 275:38067-38072(2000).
 RN [3]
 RP SEQUENCE OF 90-785 FROM N.A.; AND SEQUENCE OF 359-375 AND 670-675.
 RX MEDLINE-88080466; PubMed-3319186;
 RA Kadonaga J.T., Garner K.R., Masiazk F.R., Tjian R.;
 RT "Isolation of cDNA encoding transcription factor Sp1 and functional

RT analysis of the DNA binding domain.";
 RL Cell 51:1079-1090(1987).
 RN [4]
 RP O-GLYCOSYLATION.
 RX MEDLINE-89003041; PubMed-3139301;
 RA Jackson S.P., Tjian R.;
 RT "O-glycosylation of eukaryotic transcription factors: Implications
 RL for mechanisms of transcriptional regulation.";
 RN Cell 55:125-133(1988).
 RN [5]
 RP STRUCTURE BY NMR OF 654-684 AND 684-712.
 RX MEDLINE-97218212; PubMed-9065444;
 RA Narayan V.A., Kriwacki R.W., Caradonna J.P.;
 RT "Structures of zinc finger domains from transcription factor Sp1.
 RL Insights into sequence-specific protein-DNA recognition.";
 RN J. Biol. Chem. 272:7801-7809(1997).
 RN [6]
 RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
 RX MEDLINE-96224025; PubMed-8626793;
 RA Parks C.L., Shenk T.;
 RT "The serotonin 1a receptor gene contains a TATA-less promoter that
 RL responds to MAZ and Sp1.";
 RN J. Biol. Chem. 271:4417-4430(1996).
 CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
 CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
 CC SEROTONIN RECEPTOR PROMOTER.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF252284; AAF67726.1; -;
 DR EMBL: AB039286; BAB13476.1; -;
 DR EMBL: J031133; AAA61154.1; -;
 DR FIR; A29635; A29635.
 DR PDB; 1SP1; 21-APR-97.
 DR PDB; 1SP2; 21-APR-97.
 DR TRANSFAC; T00759; -;
 DR GlycoSuiteDB; P08047; -;
 DR MIM; 189906; -;
 DR InterPro; IPR000822; Znf-C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR SMART; SM00355; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
 FT DOMAIN 626 708
 FT ZN_FING 626 650
 FT ZN_FING 656 680
 FT ZN_FING 686 708
 FT CONFLICT 366 366
 FT CONFLICT 670 670
 FT CONFLICT 785 785
 SQ SEQUENCE 785 AA; 80693 MW; 43893DBF6518B9EA CRC64;

Query Match 91.7%; Score 22; DB 1; Length 785;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 DB 736 GTATPS 741


```

RESULT 23
SPL_RAT
ID SPL_RAT STANDARD; PRT; 788 AA.
AC Q01714;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor Spl.
GN Spl.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93010958; PubMed=1356762;
RA Imataka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,
RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
RT "Two regulatory proteins that bind to the basic transcription element
RT (BTE), a GC box sequence in the promoter region of the rat P-450IA1
RT gene.";
RL EMBO J. 11:3663-3671(1992).
CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D12768; BAA02235.1; -.
CC DR PIR; JS0747; JS0747.
CC DR HSSP; P08047; 1SP1.
CC DR TRANSFAC; T00754; -.
CC DR InterPro; IPR000822; Znf-C2H2.
CC DR PRINTS; PF00096; zf-C2H2; 3.
CC DR SMART; SM00355; Znf_C2H2; 3.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
CC DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
CC KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
CC DNA-binding; Nuclear protein; Repeat; Glycoprotein.
CC FT DOMAIN 629 711
CC FT ZN_FING 629 653 C2H2-TYPE.
CC FT ZN_FING 659 683 C2H2-TYPE.
CC FT ZN_FING 689 711 C2H2-TYPE.
CC SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;

Query Match 91.7%; Score 22; DB 1; Length 788;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
Db 739 GTATPS 744

RESULT 24
ULK1_HUMAN
ID ULK1_HUMAN STANDARD; PRT; 1050 AA.
AC O75385;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93010958; PubMed=1356762;
RA Imataka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,
RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
RT "Two regulatory proteins that bind to the basic transcription element
RT (BTE), a GC box sequence in the promoter region of the rat P-450IA1
RT gene.";
RL EMBO J. 11:3663-3671(1992).
CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF045458; AAC32326.1; -.
CC DR HSSP; P00523; 2PTK.
CC DR MIM; 603168; -.
CC DR InterPro; IPR000719; Euk_pkinase.
CC DR InterPro; IPR002290; Ser_thr_pkinase.
CC DR InterPro; IPR001245; Tyr_pkinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR SMART; SM00220; S_TKC; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC KW Transferase; Serine/threonine-protein kinase; ATP-binding.
CC FT DOMAIN 16 278
CC FT NP_BIND 22 30 ATP (BY SIMILARITY).
CC FT BINDING 46 46 ATP (BY SIMILARITY).
CC FT ACT_SITE 138 138 BY SIMILARITY.
CC FT DOMAIN 297 310 POLY-SER.
CC SQ SEQUENCE 1050 AA; 112601 MW; 4ED9B94028E3C138 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 1050;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
Db 745 GTSSPS 750

RESULT 25
GNRP_MOUSE
ID GNRP_MOUSE STANDARD; PRT; 1262 AA.
AC P27671;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Guanine nucleotide releasing protein (GNRP) (Ras-specific nucleotide
DE exchange factor CDC25) (CDC25Mm).
GN RASGRF1 OR CDC25 OR GRF1.

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=93010996; PubMed=1396590;
RA Cen H., Lowy D.B.;
RT "Isolation of multiple mouse cDNAs with coding homology to
RT Saccharomyces cerevisiae CDC25: identification of a region related to
RT Bcr, Vav, Dbl and CDC24.";
RL EMOB J. 11:4007-4015(1992).
RN [2]
RP SEQUENCE OF 791-1262 FROM N.A.
RC STRAIN=SWISS; TISSUE=Brain;
RX MEDLINE=92289680; PubMed=1376246;
RA Martegani E., Vanoni M., Zippel R., Cocchetti P., Brambilla R.,
RA Ferrari C., Sturani E.P., Alberghina L.;
RT "Cloning by functional complementation of a mouse cDNA encoding a
RT homologue of CDC25, a Saccharomyces cerevisiae RAS activator.";
RL EMOB J. 11:2151-2157(1992).
RN [3]
RP SEQUENCE OF 1031-1226 FROM N.A.
RC MEDLINE=92357779; PubMed=1379731;
RX Wei W., Mosteller R.D., Sanyal P., Gonzales E., McKinney D.,
RA Dasgupta C., Li P., Liu B.X., Broek D.;
RT "Identification of a mammalian gene structurally and functionally
RT related to the CDC25 gene of Saccharomyces cerevisiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7100-7104(1992).
CC -!- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: CONTAINS 2 PH DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
CC -!- SIMILARITY: CONTAINS 1 RASGEF DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20899; AAA02741.1; -;
DR EMBL; X59868; CAA42525.1; -;
DR PIR; S20730; S20730.
DR PIR; S22693; S22693.
DR MGI; 99694; Rasgrf1.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000651; RasGEFN.
DR InterPro; IPR001895; RasGEF_CDC25.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00169; PH; 2.
DR Pfam; PF00617; RasGEF; 1.
DR Pfam; PF00618; RasGEFN; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 2.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS00741; DH_1; 1.
DR PROSITE; PS00720; GDS_CDC25; 1.
DR PROSITE; PS50096; IQ; 1.
DR PROSITE; PS50003; PH_DOMAIN; 2.
KW Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 22 130

FT DOMAIN 208 233 IQ.
FT DOMAIN 244 430 DH.
FT DOMAIN 460 588 PH 2.
FT DOMAIN 1025 1259 RASGEF.
FT CONFLICT 1033 1033 E -> D (IN REF. 3).
SQ SEQUENCE 1262 AA; 144101 MW; 38BFE68F7C2228DC8 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 1262;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxps 6
Db 734 GTSSPS 739
|||
|||

RESULT 26
CID_DROME
ID_CID_DROME STANDARD; PRT; 1377 AA.
AC P19538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Cubitus interruptus dominant protein.
GN CI-D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=90346286; PubMed=2166702;
RA Orenic T.V., Slusarski D.C., Kroll K.L., Holmgren R.A.;
RT "Cloning and characterization of the segment polarity gene cubitus
RT interruptus dominant of Drosophila.";
RL Genes Dev. 4:1053-1067(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=92146935; PubMed=1686006;
RA Berry A.J., Ajioke J.W., Kreitman M.;
RT "Lack of polymorphism on the Drosophila fourth chromosome resulting
RT from selection.";
RL Genetics 129:1111-1117(1991).
CC -!- FUNCTION: INVOLVED IN SEGMENT POLARITY. IS REQUIRED FOR THE NORMAL
CC DEVELOPMENT OF THE POSTERIOR HALF OF EACH EMBRYONIC SEGMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED UNIFORMLY THROUGHOUT THE BLASTODERM
CC STAGE AND GASTRULATION AND DOES NOT RESOLVE INTO SEGMENTALLY
CC REPEATING STRIPES UNTIL THE END OF THE SHORT PHASE OF GERM-BAND
CC EXTENSION.
CC -!- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54360; CAA38244.1; -;
DR PIR; A35817; A35817.
DR PIR; S12769; S12769.
DR HSSP; P08151; 2GLI.
DR FlyBase; FBgn0004859; ci.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR PRINTS; PR00048; ZINCINGER.
DR

```
DR SMART; SM00355; ZnF_C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
DR DEVELOPMENTAL PROTEIN; Segmentation polarity protein; Zinc-finger;
KW Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT DOMAIN 451 603 ZINC FINGERS.
FT ZN_FING 451 476 C2H2-TYPE.
FT ZN_FING 484 511 C2H2-TYPE.
FT ZN_FING 517 541 C2H2-TYPE.
FT ZN_FING 547 572 C2H2-TYPE.
FT ZN_FING 578 603 C2H2-TYPE.
SQ SEQUENCE 1377 AA; 150881 MW; 85189314D5E7B3B7 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 1377;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 655 GTSSPS 660

RESULT 27
ZEPL_HUMAN STANDARD; PRT; 2717 AA.
AC P15822;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein 40 (human immunodeficiency virus type I enhancer-
DE binding protein 1) (HIV-Ep1) (Major histocompatibility complex binding
DE protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)
DE (PRDII-BF1).
DE HIVEP1 OR ZNF40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=90169514; PubMed=2106471;
FAO C.M., Maniatis T.;
" A DNA-binding protein containing two widely separated zinc finger
motifs that recognize the same DNA sequence.";
Genes Dev. 4:29-42(1990).
[2]
STRUCTURE BY NMR OF 2113-2142.
RX MEDLINE=91064333; PubMed=2248949;
Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,
Gronenborn A.M.;
"High-resolution three-dimensional structure of a single zinc finger
from a human enhancer binding protein in solution.";
Biochemistry 29:9324-9334(1990).
[3]
STRUCTURE BY NMR OF 2087-2142.
RX MEDLINE=92232684; PubMed=1567844;
Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,
Gronenborn A.M.;
"High-resolution solution structure of the double Cys2His2 zinc
finger from the human enhancer binding protein MBP-1.";
Biochemistry 31:3907-3917(1992).
CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
CC 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
CC IN T-CELL ACTIVATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.
CC -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
```

```
CC ZINC-FINGER IN-BETWEEN.
CC -1- SIMILARITY: STRONG, TO HIVP2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X51435; CAA35798.1; -
CC PIR; A34203; A34203.
CC PDB; 3ZNF; 15-JAN-92.
CC PDB; 4ZNF; 15-JAN-92.
CC PDB; 1BBO; 31-OCT-93.
CC TRANSFAC; T00497; -
CC MIM; 194540; -
CC InterPro; IPR000822; Znf-C2H2.
CC Pfam; PF00096; zf-C2H2; 5.
CC PRINTS; P00048; ZINCFINGER.
CC SMART; SM00355; Znf_C2H2; 4.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
KW Nuclear protein; Repeat; 3D-structure.
FT DOMAIN 406 456 ZINC FINGERS.
FT ZN_FING 406 428 C2H2-TYPE.
FT ZN_FING 434 456 C2H2-TYPE.
FT DOMAIN 803 861 POLY-SER.
FT ZN_FING 958 981 C2HC-TYPE (POTENTIAL).
FT DOMAIN 2087 2139 ZINC FINGERS.
FT ZN_FING 2087 2109 C2H2-TYPE.
FT ZN_FING 2115 2139 C2H2-TYPE.
FT STRAND 2088 2088
FT TURN 2090 2092
FT STRAND 2095 2095
FT HELIX 2099 2108
FT TURN 2109 2109
FT STRAND 2115 2116
FT STRAND 2123 2124
FT HELIX 2127 2135
SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 2717;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 185 GTTSPS 190

RESULT 28
THYG_BOVIN STANDARD; PRT; 2769 AA.
AC P01267; O18976; Q95478; Q28196;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thyroglobulin precursor.
DE TG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=85296288; PubMed=3855243;
Mercken L., Simons M.-J., Swillens S., Massaer M., Vassart G.;
"Primary structure of bovine thyroglobulin deduced from the sequence
```

of its 8,431-base complementary DNA.";
 Nature 316:647-651(1985).
 [2]
 RN SEQUENCE OF 1-930 FROM N.A.
 RP MEDLINE=85127025; PubMed=3855750;
 RX Mercken L., Simons M.-J., de Martynhoff G., Swillens S., Vassart G.;
 RA "Presence of homonogenic and repetitive domains in the first 930
 RT amino acids of bovine thyroglobulin as deduced from the CDNA
 sequence";
 RL Eur. J. Biochem. 147:59-64(1985).
 [3]
 RN SEQUENCE OF 1-22 FROM N.A.
 RP MEDLINE=87190432; PubMed=3032624;
 RX de Martynhoff G., Pohl V., Mercken L., van Ommen G.-J., Vassart G.;
 RA "Structural organization of the bovine thyroglobulin gene and of its
 RT 5'-flanking region.";
 RL Eur. J. Biochem. 164:591-599(1987).
 [4]
 RN SEQUENCE OF 1002-1209 FROM N.A.
 RP MEDLINE=88062712; PubMed=3681978; Vassart G.;
 RX Parma J., Christophe D., Pohl V.;
 RA "Structural organization of the 5' region of the thyroglobulin gene.
 RT Evidence for intron loss and 'exonization' during evolution.";
 RL J. Mol. Biol. 196:769-779(1987).
 CC -1- FUNCTION: PRECURSOR OF THE IODINATED THYROID HORMONES THYROXINE
 CC (T4) AND TRIIODOTHYRONINE (T3).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: THYROID GLAND SPECIFIC.
 CC -1- PTM: SULFATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 11 THYROGLOBULIN TYPE-I DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X02815; CAA26584.1; -;
 DR EMBL; X02155; CAA26090.1; -;
 DR EMBL; X05380; CAA28971.1; ALT_SEQ.
 DR EMBL; X06071; CAA29457.1; -;
 DR EMBL; X06072; CAA29457.1; JOINED.
 DR EMBL; X06073; CAA29457.1; JOINED.
 DR EMBL; X06074; CAA29457.1; JOINED.
 DR EMBL; X06075; CAA29457.1; JOINED.
 DR PIR; A01533; UIBO.
 DR HSSP; P21836; 1MAA.
 DR InterPro; IPR002018; Carboxylesterase_B.
 DR InterPro; IPR000716; Thyroglobulin_1.
 DR Pfam; PF00135; Coesterase; 2.
 DR Pfam; PF00086; Thyroglobulin_1; 8.
 DR SMART; SM00211; TY; 10.
 DR DR PROSITE; PS00484; THYROGLOBULIN_1; 9.
 DR DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 KW Glycoprotein; Repeat; Thyroid hormone; Iodination; Sulfation; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 2769
 FT DOMAIN 31 92 THYROGLOBULIN.
 FT DOMAIN 93 160 THYROGLOBULIN TYPE IA 1.
 FT DOMAIN 161 297 THYROGLOBULIN TYPE IA 2.
 FT DOMAIN 298 358 THYROGLOBULIN TYPE IA 3.
 FT DOMAIN 604 657 THYROGLOBULIN TYPE IA 4.
 FT DOMAIN 661 725 THYROGLOBULIN TYPE IA 5.
 FT DOMAIN 726 921 THYROGLOBULIN TYPE IA 6.
 FT DOMAIN 922 1073 THYROGLOBULIN TYPE IA 7.
 FT DOMAIN 1074 1145 THYROGLOBULIN TYPE IA 8.
 FT DOMAIN 1146 1210 THYROGLOBULIN TYPE IB 1.
 FT REPEAT 1458 1471 THYROGLOBULIN TYPE IA 9.
 FT REPEAT 1472 1488 TYPE II.

FT REPEAT 1489 1505
 FT DOMAIN 1513 1567
 FT REPEAT 1605 1725
 FT REPEAT 1726 1893
 FT REPEAT 1894 1996
 FT REPEAT 1997 2130
 FT REPEAT 2131 2188
 FT MOD_RES 24 24
 FT MOD_RES 24 24
 FT MOD_RES 2574 2574
 FT MOD_RES 2588 2588
 FT MOD_RES 2767 2767
 FT CARBOHYD 110 110
 FT CARBOHYD 198 198
 FT CARBOHYD 483 483
 FT CARBOHYD 495 495
 FT CARBOHYD 747 747
 FT CARBOHYD 853 853
 FT CARBOHYD 947 947
 FT CARBOHYD 1140 1140
 FT CARBOHYD 1365 1365
 FT CARBOHYD 1776 1776
 FT CARBOHYD 1870 1870
 FT CARBOHYD 2014 2014
 FT CARBOHYD 2123 2123
 FT CARBOHYD 2251 2251
 FT CARBOHYD 2296 2296
 FT CONFLICT 1206 1206
 SQ SEQUENCE 2769 AA; 303218 MW; 1C7F227E9101DE2A CRC64;

Query Match 91.7%; Score 22; DB 1; Length 2769;
 Best Local Similarity 66.7%; Pred. No. 5.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 |||||
 Db 2196 GTSSPS 2201

RESULT 29
 LYS3_ECOLI
 ID LYS3_ECOLI STANDARD; PRT; 45 AA.
 AC P05821;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E1 precursor.
 GN LYS.
 OS Escherichia coli.
 OG Plasmid ColE1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094231; PubMed=3936034;
 RA Waleh N.S., Johnson P.H.;
 RT "Structural and functional organization of the colicin E1 operon";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:8389-8393(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC31, AND EC71;
 RX MEDLINE=95062249; PubMed=7972047;
 RA Riley M.A., Tan Y., Wang J.;
 RT "Nucleotide polymorphism in colicin E1 and Ia plasmids from natural
 RT isolates of Escherichia coli";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11276-11280(1994).
 CC -1- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 CC AND PARTIAL CELL LYSIS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M12543; AAA23067.1; -;
 CC EMBL; U15629; AAA59411.1; -;
 CC EMBL; U15633; AAA59419.1; -;
 CC PIR; B24685; ZHECP1.
 CC InterPro; IPR003059; Lysis_col.
 CC Pfam; PF02402; Lysis_col; 1.
 CC PRINTS; PR01297; LYSISCOLICIN.
 CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Plasmid; Outer membrane; Lipoprotein; Signal.
 CC SIGNAL 1 17 POTENTIAL.
 CC FT CHAIN 18 45 LYSIS PROTEIN FOR COLICIN E1.
 CC FT LIPID 18 18 N-ACYL DIGLYCERIDE (POTENTIAL).
 CC SEQUENCE 45 AA; 4829 MW; A34D3B4FBC12A13E CRC64;

Query Match 87.5%; Score 21; DB 1; Length 45;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY 1 gtxxps 6
 ||||
 Db 29 GTTAPS 34

RESULT 30

LYS3_SHISO STANDARD; PRT; 45 AA.
 AC P21185;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E1* precursor.
 GN KIL.
 OS Shigella sonnei.
 OG Plasmid pKY-1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Shigella.
 OC NCBI_TaxID=624;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Higashi M., Hata M., Hase T., Yamaguchi K., Masamune Y.;
 RT "The nucleotide sequence of cea and the region of origin of plasmid
 RT pKY-1.";
 RL J. Gen. Appl. Microbiol. 32:433-442(1986).
 CC -1- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 CC AND PARTIAL CELL LYSIS.
 CC -1- SIMILARITY: 96% IDENTITY TO E. COLI LYSIS PROTEIN PRECURSOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M37218; AAA98158.1; -;
 CC PIR; S10921.
 CC InterPro; IPR003059; Lysis_col.
 CC Pfam; PF02402; Lysis_col; 1.
 CC PRINTS; PR01297; LYSISCOLICIN.
 CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Plasmid; Outer membrane; Lipoprotein; Signal.
 CC SIGNAL 1 17 POTENTIAL.
 CC FT CHAIN 18 45 LYSIS PROTEIN FOR COLICIN E1*.
 CC FT LIPID 18 18 N-ACYL DIGLYCERIDE (POTENTIAL).
 CC SEQUENCE 45 AA; 4831 MW; A34D232FBC12B99D CRC64;

Query Match 87.5%; Score 21; DB 1; Length 45;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY 1 gtxxps 6
 ||||
 Db 29 GTVAPS 34

RESULT 31

LYS2_ECOLI STANDARD; PRT; 47 AA.
 AC P06963;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicins E2 and E3 precursor.
 GN HIC OR CELB.
 OS Escherichia coli.
 OG Plasmid ColE2-P9, and Plasmid ColE3-CA38.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Plasmid-ColE2-P9;
 RX MEDLINE=86195936; PubMed=3516985;
 RA Toba M., Masaki H., Ohta T.;
 RT "Primary structures of the ColE2-P9 and ColE3-CA38 lysis genes.";
 RL J. Biochem. 99:591-596(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Plasmid-ColE2-P9;
 RX MEDLINE=85239907; PubMed=3892228;
 RA Cole S.T., Saint-Joanis B., Pugsley A.P.;
 RT "Molecular characterisation of the colicin E2 operon and
 RT identification of its products.";
 RL Mol. Gen. Genet. 198:465-472(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Plasmid-ColE3-CA38;
 RX MEDLINE=85028427; PubMed=6092219;
 RA Watson R.J., Lau P.C.K., Vernet T., Visentin L.P.;
 RT "Characterization and nucleotide sequence of a colicin-release gene
 RT in the hic region of plasmid ColE3-CA38.";
 RL Gene 29:175-184(1984).
 RN [4]
 RP ERRATUM.
 RC PLASMID-ColE3-CA38;
 RA Watson R.J., Lau P.C.K., Vernet T., Visentin L.P.;
 RL Gene 42:351-353(1986).
 RN [5]
 RP SEQUENCE OF 1-38 FROM N.A.
 RC PLASMID-ColE3-CA38;
 RX MEDLINE=85210906; PubMed=3889348;
 RA Masaki H., Ohta T.;
 RT "Colicin E3 and its immunity genes.";
 RL J. Mol. Biol. 182:217-227(1985).
 CC -1- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 CC AND PARTIAL CELL LYSIS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X02397; CAA26243.1; -;
 CC EMBL; D00020; BAA00014.1; -;

DR EMBL; D00021; BAA00015.1; -;
 DR EMBL; J01574; AA88419.1; -;
 DR EMBL; X03631; CAA27281.1; -;
 DR EMBL; X03632; CAA27282.1; -;
 DR EMBL; M29885; AAA23070.1; -;
 DR PIR; A22383; BVECH8.
 DR PIR; JS0004; JS0004.
 DR PIR; S09529; S09529.
 DR InterPro; IPR003059; Lysis_col.
 DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISCOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Plasmid; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 FT SEQUENCE 47 AA; 4860 MW; 691E149A8A164A0A CRC64;
 SQ

Query Match 87.5%; Score 21; DB 1; Length 47;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 DB 31 GTVSPS 36

RESULT 32
 LYS5_ECOLI
 ID LYS5_ECOLI STANDARD; PRT; 47 AA.
 AC PI3344;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E5 precursor.
 GN LYS.
 OS Escherichia coli.
 OG Plasmid ColE5-099.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93364708; PubMed=2549375;
 RA Lau P.C.K., Condie J.A.;
 RT "Nucleotide sequences from the colicin E5, E6 and E9 operons:
 presence of a degenerate transposon-like structure in the ColE9-J
 plasmid.";
 RT Mol. Gen. Genet. 217:269-277(1989).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90218006; PubMed=2561131;
 RA Curtis M.D., James R., Coddington A.;
 RT "An evolutionary relationship between the ColE5-099 and the ColE9-J
 plasmids revealed by nucleotide sequencing.";
 RL J. Gen. Microbiol. 135:2783-2788(1989).
 CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 AND PARTIAL CELL LYSIS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X15857; CAA33861.1; -;
 DR EMBL; M30445; AAA98053.1; -;
 DR PIR; JQ0330; JQ0330.
 DR PIR; C45799; C45799.
 DR InterPro; IPR003059; Lysis_col.

DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISCOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Plasmid; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 FT SEQUENCE 47 AA; 4926 MW; 69133B25CA15A4B8 CRC64;
 SQ

Query Match 87.5%; Score 21; DB 1; Length 47;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 DB 31 GTVSPS 36

RESULT 33
 LYS6_ECOLI
 ID LYS6_ECOLI STANDARD; PRT; 47 AA.
 AC PI3345;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E6 precursor.
 GN LYS.
 OS Escherichia coli.
 OG Plasmid ColE6-CT14.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89364708; PubMed=2549375;
 RA Lau P.C.K., Condie J.A.;
 RT "Nucleotide sequences from the colicin E5, E6 and E9 operons:
 presence of a degenerate transposon-like structure in the ColE9-J
 plasmid.";
 RT Mol. Gen. Genet. 217:269-277(1989).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90078082; PubMed=2687234;
 RA Akutsu A., Masaki H., Ohta T.;
 RT "Molecular structure and immunity specificity of colicin E6, an
 evolutionary intermediate between E-group colicins and cloacin
 Df13.";
 RL J. Bacteriol. 171:6430-6436(1989).
 CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 AND PARTIAL CELL LYSIS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X15856; CAA33858.1; -;
 DR EMBL; M31808; AAA23083.1; -;
 DR PIR; JQ0328; JQ0328.
 DR PIR; D43716; D43716.
 DR InterPro; IPR003059; Lysis_col.
 DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISCOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Plasmid; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 FT SEQUENCE 47 AA; 4902 MW; 691E124F6A057ED8 CRC64;
 SQ

Query Match 87.5%; Score 21; DB 1; Length 47;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| || ||
Db 31 GTVSPS 36

RESULT 34

LYS7_ECOLI STANDARD; PRT; 47 AA.
AC Q03709;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysis protein for colicin E7 precursor.
GN LYS OR CELE7.
OS Escherichia coli.
OC Plasmid Cole7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91259043; PubMed=2045785;
RA Chak K.F., Kuo W.S., Lu F.M., James R.;
RT "Cloning and characterization of the Cole7 plasmid.";
RL J. Gen. Microbiol. 137:91-100(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K317;
RA Lau P.C.K., Parsons M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
CC AND PARTIAL CELL LYSIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M57540; AAA23072.1; -;
DR EMBL: X63620; CAA45166.1; -;
DR InterPro: IPR003059; Lysis_col.
DR Pfam: PF02402; Lysis_col; 1.
DR PRINTS: PR01297; LYSISCOLICIN.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Plasmid; Outer membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 47 LYSIS PROTEIN FOR COLICIN E7.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 47 AA; 4874 MW; 6900F2A17A057ED8 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 47;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| || ||
Db 31 GTVSPS 36

RESULT 35

LYS8_ECOLI STANDARD; PRT; 47 AA.
ID LYS8_ECOLI
AC P10099;

DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysis protein for colicin E8 precursor.
GN LYS.
OS Escherichia coli.
OC Plasmid Cole8.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88121677; PubMed=3323826;
RA Uchimura T., Lau P.C.K.;
RT "Nucleotide sequences from the colicin E8 operon: homology with
RT plasmid Cole2-P9.";
RL Mol. Gen. Genet. 209:489-493(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257046; PubMed=3290201;
RA Tobia M., Masaki H., Ohta T.;
RT "Colicin E8, a DNase which indicates an evolutionary relationship
RT between colicins E2 and E3.";
RL J. Bacteriol. 170:3237-3242(1988).
CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
CC AND PARTIAL CELL LYSIS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M21404; AAA23075.1; -;
DR EMBL: X06119; CAA29493.1; -;
DR PIR: C28184; ZHECE8.
DR PIR: S01082; S01082.
DR InterPro: IPR003059; Lysis_col.
DR Pfam: PF02402; Lysis_col; 1.
DR PRINTS: PR01297; LYSISCOLICIN.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Plasmid; Outer membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 47 LYSIS PROTEIN FOR COLICIN E8.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 47 AA; 4844 MW; 691E149A8A056B38 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 47;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| || ||
Db 31 GTVSPS 36

RESULT 36

LYS9_ECOLI STANDARD; PRT; 47 AA.
ID LYS9_ECOLI
AC P15176;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysis protein for colicin E9 precursor.
GN LYS.
OS Escherichia coli.
OC Plasmid Cole9-J.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;

```

RN RP SEQUENCE FROM N.A.
RA Lau P.C.K.;
RL Submitted (DCC-1989) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88034907; PubMed=3312476;
RA James R., Jarvis M., Barker D.F.;
RT "Nucleotide sequence of the immunity and lysis region of the Cole9-J
plasmid."
RL J. Gen. Microbiol. 133:1553-1562(1987).
CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
CC AND PARTIAL CELL LYSIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15858; CAA33866.1; -
DR EMBL; M16803; AAA23079.1; -
DR PIR; D32535; D32535.
DR InterPro: IPR003059; Lysis_col.
DR Pfam; PF02402; Lysis_col; 1.
DR PRINTS; PR01297; LYSISCOLICIN.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Plasmid; Outer membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 47
FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 47 AA; 4816 MW; 69062F749A057ED8 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 47;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 31 GTVSPS 36

RESULT 37
LYS4_ECOLI STANDARD; PRT; 49 AA.
AC P02987;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysis protein precursor (Protein H).
GN H OR CEX.
OS Escherichia coli.
OG Plasmid Clo DF13.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86314306; PubMed=3749334;
RA Nijkamp H.J.J., de Lang R., Stuitje A.R., van den Elsen P.J.M.,
RA Veltkamp E., van Putten A.J.;
RT "The complete nucleotide sequence of the bacteriocinogenic plasmid
CloDF13."
RL Plasmid 16:135-160(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81148852; PubMed=6163089;
RA Stuitje A.R., Spelt C.E., Veltkamp E., Nijkamp H.J.J.;
RT "Identification of mutations affecting replication control of plasmid
Clo DF13."

```

```

RL Nature 290:264-267(1981).
CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
CC AND PARTIAL CELL LYSIS.
CC -!- MISCELLANEOUS: PLASMID CLO DF13 ORIGINATES FROM ENTEROBACTER
CC CLOACAE BUT IS STABLY MAINTAINED IN AND STUDIED MOSTLY FROM
CC E.COLI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04466; CAA28145.1; -
DR PIR; A03515; ZHECP3.
DR InterPro: IPR003059; Lysis_col.
DR Pfam; PF02402; Lysis_col; 1.
DR PRINTS; PR01297; LYSISCOLICIN.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Plasmid; Outer membrane; Lipoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 49
FT LIPID 22 22 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 49 AA; 5157 MW; 1A3DC979EBB0C5DC CRC64;

Query Match 87.5%; Score 21; DB 1; Length 49;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 33 GTVAPS 38

RESULT 38
LYS4_ECOLI STANDARD; PRT; 52 AA.
AC P09181;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysis protein for colicin N precursor.
GN CNL.
OS Escherichia coli.
OG Plasmid CloN pCHAP4.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-K12;
RX MEDLINE=88174431; PubMed=3280946;
RA Pugsley A.P.;
RT "The immunity and lysis genes of CloN plasmid pCHAP4."
RL Mol. Gen. Genet. 211:335-341(1988).
CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
CC AND PARTIAL CELL LYSIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06933; CAA30019.1; -
DR PIR; S01760; ZHECN4.
DR InterPro: IPR003059; Lysis_col.
DR Pfam; PF02402; Lysis_col; 1.

```


DR PRINTS: PR01297; LYSISCOLICIN.
 DR PROSITE; PS00013; PROXAR-LIPOPROTEIN; 1.
 KW Plasmid; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 52
 FT LIPID 18 18 N-ACYL DIGLYCERIDE.
 SQ SEQUENCE 52 AA; 5633 MW; F6F1BF9BDAF81B6 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 52;
 Best Local Similarity 66.7%; Pred. No. 26;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||
 DB 29 GTVAPS 34

RESULT 39
 PAHO_CHICK STANDARD; PRT; 80 AA.
 ID PAHO_CHICK
 AC POL306;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pancreatic hormone precursor (Pancreatic polypeptide) (PP).
 OS Gallus gallus (Chicken), and
 OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031, 9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Chicken; TISSUE=Liver;
 RX MEDLINE=93366173; PubMed=8359635;
 RA Nata K., Sugimoto T., Kohri K., Hidaka H., Hattori E., Yamamoto H.,
 RA Yonekura H., Okamoto H.;
 RT "Structure determination and evolution of the chicken cDNA and gene
 RT encoding prepancreatic polypeptide.";
 RL Gene 130:183-189(1993).
 RN [2]
 RP SEQUENCE OF 26-61.
 RC SPECIES=Chicken;
 RX MEDLINE=76069270; PubMed=1194289;
 RA Kimmel J.R., Hayden L.J., Pollock H.G.;
 RT "Isolation and characterization of a new pancreatic polypeptide
 RT hormone.";
 RL J. Biol. Chem. 250:9369-9376(1975).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RC SPECIES=M.gallopavo;
 RX MEDLINE=84179397; PubMed=6673760;
 RA Glover I., Maneef I., Pitts J., Woods S., Moss D., Tickle I.,
 RA Blundell T.L.;
 RT "Conformational flexibility in a small globular hormone: X-ray
 RT analysis of avian pancreatic polypeptide at 0.98-A resolution.";
 RL Biopolymers 22:293-304(1983).
 CC -!- FUNCTION: PANCREATIC HORMONE IS SYNTHESIZED IN PANCREATIC ISLETS
 CC OF LANGERHANS AND ACTS AS A REGULATOR OF PANCREATIC AND
 CC GASTROINTESTINAL FUNCTIONS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D13761; BAA02907.1; -.

DR EMBL; D13760; BAA02906.1; -.
 DR PIR; A01575; PCCH.
 DR PDB; 1PPT; 15-OCT-91.
 DR InterPro; IPR001955; Pancreatic_hormn.
 DR Pfam; PF00159; hormone3; 1.
 DR PRINTS; PR00278; PANCHORMONE.
 DR ProDom; PD001267; Pancreatic_hormn; 1.
 DR SMART; SM00309; PAH; 1.
 DR PROSITE; PS00285; PANCREATIC_HORMONE_1; 1.
 DR PROSITE; PS00276; PANCREATIC_HORMONE_2; 1.
 KW Hormone; Cleavage on pair of basic residues; Pancreas; Signal;
 KW Amidation; 3D-structure.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 61 PANCREATIC HORMONE.
 FT MOD_RES 61 61 AMIDATION (G-62 PROVIDE AMIDE GROUP).
 FT CONFLICT 47 48 ND -> DN (IN REF. 2).
 FT TURN 35 36
 FT HELIX 39 56
 FT TURN 57 58
 SQ SEQUENCE 80 AA; 8773 MW; 90B44E27389DB050 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 80;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||
 DB 23 GTAGPS 28

RESULT 40
 YQGV_BACSU STANDARD; PRT; 94 AA.
 ID YQGV_BACSU
 AC P54499;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 10.5 kDa protein in SODA-COMGA intergenic region.
 GN YQGV.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,
 RA Sato T., Takeuchi M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE UPF0045 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D84432; BAA12525.1; -.
 DR EMBL; Z99116; CAB14412.1; -.
 DR Subtilisin; BGL1689; YQGV.
 DR InterPro; IPR002767; DUF77.
 DR Pfam; PF01910; DUF77; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 94 AA; 10453 MW; 4A576F2338B80999 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 94;
 Best Local Similarity 66.7%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 gtxxps 6
DB 12 GTEPS 17

RESULT 41
TRH4_ECOLI
ID TRH4_ECOLI STANDARD; PRT; 118 AA.
AC Q00190;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE TRAH protein.
GN TRAH.
OS Escherichia coli.
OG Plasmid IncP-beta RP4.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.
RC STRAIN-HB101;
RX MEDLINE=92190548; PubMed=1665997;
RA Ziegelin G., Pansegrau W., Strack B., Balzer D., Kroeger M.,
RA Kruff V., Lanka E.;
RT "Nucleotide sequence and organization of genes flanking the transfer
RL DNA Seq. 1:303-327(1991).
CC -!- FUNCTION: THE INITIATION PROCESS OF TRANSFER DNA SYNTHESIS
CC REQUIRES THE INTERACTION OF AT LEAST THREE PLASMID-SPECIFIC
CC COMPONENTS (TRAH, I, AND J) AT THE TRANSFER ORIGIN RESULTING
CC IN THE ASSEMBLY OF A SPECIALISED NUCLEOPROTEIN COMPLEX - THE
CC RELAXOSOME.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54459; CAA38335.1; -.
CC PIR; S23000; S23000.
CC Plasmid; Conjugation.
FT INIT_MET 0
SQ SEQUENCE 118 AA; 12738 MW; 2FC1445479E6F4B6 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 118;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 37 GTLAPS 42

RESULT 42
RK14_PORPU
ID RK14_PORPU STANDARD; PRT; 122 AA.
AC P51304;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chloroplast 50S ribosomal protein L14.
GN RPL14.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 87.5%; Score 21; DB 1; Length 118;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 37 GTLAPS 42

RESULT 43
FMF7_ECOLI
ID FMF7_ECOLI STANDARD; PRT; 170 AA.
AC P25394;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE F107 fimbrial protein precursor.
GN FEDE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=107/86;
RX MEDLINE=92225617; PubMed=1348723;
RA Imberechts H.A., de Greve H., Schlicker C., Bouchet H., Pohl P.,
RA Charlier G., Vandekerckhove J., van Damme J., van Montagu M.,
RA Lintermans P.;
RT "Characterization of F107 fimbriae of Escherichia coli 107/86, which
RT causes edema disease in pigs, and nucleotide sequence of the F107
RT major fimbrial subunit gene, feda.";
RL Infect. Immun. 60:1963-1971(1992).
CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC -!- SUBCELLULAR LOCATION: Fimbria.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38804; AAC08190.1; -.
CC HSP; P04450; IWHI.
DR Mendel; 10314; PORPU; rpl14; 1.
DR InterPro; IPR000218; Ribosomal_L14.
DR Pfam; PF00238; Ribosomal_L14; 1.
DR ProDom; PD001093; Ribosomal_L14; 1.
DR PROSITE; PS00049; RIBOSOMAL_L14; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 122 AA; 13411 MW; 49892C2AB0B6BEB2 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 122;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 26 GTSNPS 31

```

```

DR EMBL; M61713; AAA23735.1; -
DR EMBL; M61713; AAA23734.1; -
DR PIR; A43841; A43841.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Fimbrilia; Signal.
FT SIGNAL 1 21 F107 FIMBRIAL PROTEIN.
FT CHAIN 22 170 PROBABLE.
FT DISULFID 37 78
SQ SEQUENCE 170 AA; 17329 MW; AF40947CE387692F CRC64;

Query Match 87.5%; Score 21; DB 1; Length 170;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   || ||
Db 54 GTVAPS 59

RESULT 44
OM24_ARATH STANDARD; PRT; 187 AA.
AC P82805;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Mitochondrial import receptor subunit TOM20-4 (Translocase of outer
DE membrane 20 kDa subunit 4).
GN TOM20-4 OR AT5G40930 OR MWG1.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RA STRAIN-CV, COLUMBIA;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki A., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yokum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirkse W., Molijman P., Klein Lankhorst R.,
RA Weitzenecker T., Bothe G., Rose M., Hauf J., Berneiser S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer R.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
thaliana."
RL Nature 408:823-826(2000).
[2]
SEQUENCE OF 102-117.
RC STRAIN-CV, COLUMBIA;
RX PubMed=11161051;
RA Werhahn W., Niemeyer A., Jaensch L., Kruft V., Schmitz U.K.,
RA Braun H.-P.;

```

```

RT "Purification and characterization of the preprotein translocase of
RT the outer mitochondrial membrane from Arabidopsis thaliana.
RT Identification of multiple forms of TOM20."
RL Plant Physiol. 125:943-954(2001).
CC -|- FUNCTION: CENTRAL COMPONENT OF THE RECEPTOR COMPLEX RESPONSIBLE
CC FOR THE RECOGNITION AND TRANSLLOCATION OF CYTOSOLICALLY SYNTHESIZED
CC MITOCHONDRIAL PREPROTEINS. TOGETHER WITH TOM22 FUNCTIONS AS THE
CC TRANSIT PEPTIDE RECEPTOR AT THE SURFACE OF THE MITOCHONDRION
CC OUTER MEMBRANE AND FACILITATES THE MOVEMENT OF PREPROTEINS INTO
CC THE TRANSLLOCATION PORE.
CC -|- SUBUNIT: FORMS PART OF MITOCHONDRIAL RECEPTOR COMPLEX.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane.
CC -|- PTM: THE N-TERMINUS IS BLOCKED.
CC -|- MISCELLANEOUS: THERE ARE FOUR GENES (TOM20-1, TOM20-2, TOM20-3 AND
CC TOM20-4) WHICH ENCODE MITOCHONDRIAL IMPORT RECEPTOR SUBUNITS
CC TOM20.
CC -|- SIMILARITY: BELONGS TO THE TOM20 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB023040; BAB10523.1; -
KW Transport; Protein transport; Outer membrane; Mitochondrion;
KW Transmembrane.
FT DOMAIN 1 160 INTERMEMBRANE (POTENTIAL).
FT TRANSMEM 161 178 POTENTIAL.
FT DOMAIN 179 187 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 187 AA; 20973 MW; 156DF3D231EA286C CRC64;

Query Match 87.5%; Score 21; DB 1; Length 187;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   || ||
Db 137 GTAGPS 142

RESULT 45
Y319_MYCPN STANDARD; PRT; 193 AA.
AC P75329;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG319 homolog (H08_orf193).
GN MFN454 OR MF387.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
[1]
SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Hammelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE000037; AAB96035.1; -
 DR Hypothetical protein; Transmembrane; Complete proteome.
 KW TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 SQ SEQUENCE 193 AA; 21443 MW; BAE6806C6D80CID CRC64;

Query Match 87.5%; Score 21; DB 1; Length 193;
 Best Local Similarity 66.7%; Pred. No. 91;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 31 GTVSPS 36

RESULT 46
 COX3_MYCTU STANDARD; PRT; 203 AA.
 ID COX3_MYCTU
 AC Q10385;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cytochrome c oxidase polypeptide III (EC 1.9.3.1) (Cytochrome
 AA3 subunit 3).
 DE CTA6 OR RV2193 OR MT2249 OR MTCY190.04.
 GN Mycobacterium tuberculosis.
 OS Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; 270283; CAA94262.1; -

DR EMBL; AE007071; AAK46535.1; -
 DR HSSP; P00415; 10CC.
 DR TIGR; MT2249; -
 DR Tuberculin; RV2193; -
 DR InterPro: IPR000298; CytC_oxdase_III.
 DR Pfam; PF00510; COX3; 1.
 DR PROSITE; PS02553; COX3; 1.
 KW Oxidoreductase; Transmembrane; Complete proteome.
 FT TRANSMEM 30 50 POTENTIAL.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 96 116 POTENTIAL.
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 SQ SEQUENCE 203 AA; 22420 MW; 70145380A05BD0C2 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 203;
 Best Local Similarity 66.7%; Pred. No. 95;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 128 GTSIPS 133

RESULT 47
 NODE_RHISN STANDARD; PRT; 215 AA.
 ID NODE_RHISN
 AC P50355;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Chitooligosaccharide deacetylase (EC 3.5.1.-) (Nodulation protein B).
 GN NODB OR Y4HH.
 OS Rhizobium sp. (strain NGR234).
 OG Plasmid sym pNGR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RA "Molecular basis of symbiosis between Rhizobium and legumes."
 RT Nature 387:394-401(1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=95075295; PubMed=7984092;
 RX Relic B., Perret X., Estrada-Garcia M.T., Kopicinska J., Golinowski W.,
 RA Krishnan H.B., Pueppke S.G., Broughton W.J.;
 RT "Nod factors of Rhizobium are a key to the legume door."
 RL Mol. Microbiol. 13:171-178(1994).
 CC -1- FUNCTION: IS INVOLVED IN GENERATING A SMALL HEAT-STABLE COMPOUND
 CC (NOD), AN ACYLATED OLIGOMER OF N-ACETYLGLUCOSAMINE, THAT
 CC STIMULATES MITOSIS IN VARIOUS PLANT PROTOPLASTS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: TO OTHER POLYSACCHARIDE DEACETYLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; X73362; CAA51773.1; -

EMBL; A5000076; AAB91696.1; -

DR InterPro: IPR002509; Polysac_deacet.

DR Pfam; PF01522; Polysac_deacet; 1.

KW Hydrolase; Nodulation; Plasmid.

FT CONFLICT 148 149 RP -> S (IN REF. 2).

```

SQ SEQUENCE 215 AA; 23625 MW; 8A91E2B4E9D6D57F CRC64;

Query Match 87.5%; Score 21; DB 1; Length 215;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
  || ||
DB 14 GTGAPS 19

RESULT 48
ID XYN2_ASPNG STANDARD; PRT; 225 AA.
AC P55330; Q12557;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Endo-1,4-beta-xylanase II precursor (EC 3.2.1.8) (Xylanase II)
DE (1,4-beta-D-xylan xylanohydrolase II).
GN XYNB.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; microsporid Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 4066;
RA Ito K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -!- PATHWAY: XLAN DEGRADATION.
CC -!- SIMILARITY: BELONGS TO CELLULASE FAMILY G (FAMILY 11 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isp-sib.ch).
CC -----
DR EMBL; D38071; BAA07265.1; -
DR HSSP; P09850; 1XNB.
DR InterPro; IPR001137; Glyco_hydro_11.
DR Pfam; PF00457; Glyco_hydro_11; 1.
DR PRINTS; PR00911; GLHYDRLASE11.
DR PROSITE; PS00776; GLYCOSYL_HYDROL_F11_1; 1.
DR PROSITE; PS00777; GLYCOSYL_HYDROL_F11_2; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 37
FT CHAIN 38 225 ENDO-1,4-BETA-XYLANASE II.
FT ACT_SITE 121 121 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 212 212 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 225 AA; 24057 MW; C4B8BB007AB2B8FD CRC64;

Query Match 87.5%; Score 21; DB 1; Length 225;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
  || ||
DB 99 GTTPPS 104

RESULT 49
ID PDXJ_NEIMA STANDARD; PRT; 242 AA.

```

```

AC Q9ROV9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Pyridoxal phosphate biosynthetic protein pdxJ.
GN PDXJ OR NMA2037.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RA Linz B.; Schenker M.; Achtman M.;
RT "Frequent horizontal genetic exchange between Neisseria meningitidis
RL and commensal neisseriae.";
RN Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J.; Achtman M.; James K.D.; Bentley S.D.; Churcher C.;
RA Klee S.R.; Morelli G.; Basham D.; Brown D.; Chillingworth T.;
RA Davies R.M.; Davis P.; Devlin K.; Feltwell T.; Hamlin N.; Holroyd S.;
RA Jagels K.; Leather S.; Moule S.; Mungall K.; Quail M.A.;
RA Rajadream M.A.; Rutherford K.M.; Simmonds M.; Skelton J.;
RA Whitehead S.; Spratt B.G.; Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RL meningitidis Z2491.";
RL Nature 404:502-506(2000).
CC -!- FUNCTION: CATALYZES THE CONDENSATION OF 1-DEOXY-D-XYLULOSE-5-
CC PHOSPHATE (DXP) AND 1-AMINO-3-OXO-4-(PHOSPHOHYDROXY)PROPAN-2-ONE
CC TO FORM PYRIDOXINE 5'-PHOSPHATE (PNP) (BY SIMILARITY).
CC -!- PATHWAY: DE NOVO SYNTHESIS OF PYRIDOXINE (VITAMIN B6) AND
CC PYRIDOXAL PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE PDXJ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isp-sib.ch).
CC -----
DR EMBL; AF058689; AAF06688.1; -
DR EMBL; AL162757; CAB85256.1; -
KW Pyridoxine biosynthesis; Complete proteome.
SQ SEQUENCE 242 AA; 26669 MW; 07AC9E95DDA8D090 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 242;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
  || ||
DB 18 GTTPPS 23

RESULT 50
PDXJ_NEIMB STANDARD; PRT; 242 AA.
AC Q9KOV9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyridoxal phosphate biosynthetic protein pdxJ.
GN PDXJ OR NMB0448.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=201755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: CATALYZES THE CONDENSATION OF 1-DEOXY-D-XYLULOSE-5-
CC PHOSPHATE (DXP) AND 1-AMINO-3-OXO-4-(PHOSPHOHYDROXY)PROPAN-2-ONE
CC TO FORM PYRIDOXINE 5'-PHOSPHATE (PNP) (BY SIMILARITY).
CC -1- PATHWAY: DE NOVO SYNTHESIS OF PYRIDOXINE (VITAMIN B6) AND
CC PYRIDOXAL PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE PDJX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF002401; AAF40885.1; -.
CC TIGR; NMB0448; -.
KW Pyridoxine biosynthesis; Complete proteome.
SQ SEQUENCE 242 AA; 26565 MW; 5DA0476728AA1485 CRC64;

```

```

Query Match      87.5%; Score 21; DB 1; Length 242;
Best Local Similarity 66.7%; Pred. No. 1.le+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 gtxxps 6
   |||
Db 18 GTTYP5 23

```

```

Search completed: September 24, 2002, 11:27:08
Job time: 372 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:26:14 ; Search time 104.01 Seconds
(without alignments)
9.980 Million cell updates/sec

Title: BASK-853-CLAIM5
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

- Database :
- 1: sp_remb19:*
 - 2: sp_archaea:*
 - 3: sp_bacteria:*
 - 4: sp_fungi:*
 - 5: sp_human:*
 - 6: sp_invertebrate:*
 - 7: sp_mammal:*
 - 8: sp_mhc:*
 - 9: sp_organelle:*
 - 10: sp_phase:*
 - 11: sp_plant:*
 - 12: sp_rodent:*
 - 13: sp_virus:*
 - 14: sp_vertebrate:*
 - 15: sp_unclassified:*
 - 16: sp_virus:*
 - 17: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	71	15 Q9IF05	Q9IF05 human immun
2	22	91.7	75	10 Q41118	Q41118 phaseolus v
3	22	91.7	80	15 Q90QH8	Q90QH8 human immun
4	22	91.7	89	16 Q9HW24	Q9HW24 pseudomonas
5	22	91.7	92	15 Q38158	Q38158 human immun
6	22	91.7	93	15 Q38175	Q38175 human immun
7	22	91.7	93	15 Q38177	Q38177 human immun
8	22	91.7	96	15 Q38114	Q38114 human immun
9	22	91.7	96	15 Q38161	Q38161 human immun
10	22	91.7	96	15 Q38178	Q38178 human immun
11	22	91.7	99	15 Q38030	Q38030 human immun
12	22	91.7	99	15 Q38362	Q38362 human immun
13	22	91.7	99	15 Q38371	Q38371 human immun
14	22	91.7	99	15 Q38372	Q38372 human immun
15	22	91.7	101	15 Q38113	Q38113 human immun
16	22	91.7	101	15 Q38128	Q38128 human immun

17	22	91.7	101	15 Q38130	Q38130 human immun
18	22	91.7	101	15 Q38133	Q38133 human immun
19	22	91.7	101	15 Q38138	Q38138 human immun
20	22	91.7	101	15 Q38141	Q38141 human immun
21	22	91.7	105	17 Q9YA39	Q9YA39 aeropyrum p
22	22	91.7	111	15 Q38165	Q38165 human immun
23	22	91.7	111	15 Q38166	Q38166 human immun
24	22	91.7	111	15 Q38169	Q38169 human immun
25	22	91.7	111	15 Q38171	Q38171 human immun
26	22	91.7	112	15 Q75698	Q75698 human immun
27	22	91.7	135	11 Q64053	Q64053 mus sp. orf
28	22	91.7	142	11 Q9D5C6	Q9D5C6 mus musculus
29	22	91.7	156	4 Q9HBN7	Q9HBN7 homo sapien
30	22	91.7	157	4 Q9NWG3	Q9NWG3 homo sapien
31	22	91.7	175	5 Q9VSG7	Q9VSG7 drosophila
32	22	91.7	177	12 Q41995	Q41995 maize rayad
33	22	91.7	178	2 Q9KY82	Q9KY82 streptomyce
34	22	91.7	182	5 Q9VSO6	Q9VSO6 drosophila
35	22	91.7	186	4 Q9B294	Q9B294 homo sapien
36	22	91.7	196	10 Q82664	Q82664 arabidopsis
37	22	91.7	200	2 Q9F5I3	Q9F5I3 agrobacteri
38	22	91.7	202	16 Q9KQL1	Q9KQL1 vibrio chol
39	22	91.7	218	11 Q9JLH5	Q9JLH5 rattus norv
40	22	91.7	228	2 Q9RDK2	Q9RDK2 streptomyce
41	22	91.7	233	2 Q9KXV0	Q9KXV0 streptomyce
42	22	91.7	237	6 P79289	P79289 sus scrofa
43	22	91.7	247	2 Q9L203	Q9L203 streptomyce
44	22	91.7	253	2 Q9LJ26	Q9LJ26 streptomyce
45	22	91.7	253	4 Q43563	Q43563 homo sapien
46	22	91.7	254	16 Q9I170	Q9I170 pseudomonas
47	22	91.7	275	13 Q98TY7	Q98TY7 brachydanio
48	22	91.7	285	10 Q9MOA9	Q9MOA9 arabidopsis
49	22	91.7	288	2 Q9RNC9	Q9RNC9 bartonella
50	22	91.7	288	12 Q9DWD6	Q9DWD6 rat cytomeg

ALIGNMENTS

RESULT 1

Q9IF05	PRELIMINARY;	PRT;	71 AA.
ID	Q9IF05		
AC	Q9IF05		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	GAG PROTEIN (FRAGMENT).		
GN	POL.		
OS	Human immunodeficiency virus type 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=1079;		
RX	MEDLINE=21443958; PubMed=11559796;		
RA	Peters S., Manoz M., Yerly S., Sanchez-Merino V., Lopez-Galindez C.,		
RA	Perrin L., Larder B., Cmarko D., Fakan S., Meylan P., Teienti A.;		
RT	"Resistance to Nucleoside Analog Reverse Transcriptase Inhibitors		
RL	Mediated by Human Immunodeficiency Virus Type 1 p6 Protein.";		
DR	J. Virol. 75:9644-9653(2001).		
EMBL	AF282969; AAF87830.1; -		
FT	NON_TER 1 1		
SQ	SEQUENCE 71 AA: 7870 MW; 033C600C7F7FBCF3 CRC64;		

Query Match 91.7%; Score 22; DB 15; Length 71;

Best Local Similarity 66.7%; Pred. No. 1e+02; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 2;

Qy 1 gtxxps 6

Db 39 GTATPS 44

```

RESULT 2
Q41118 ID Q41118 PRELIMINARY; PRT; 75 AA.
AC Q41118;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHALCONE SYNTHASE MRNA (FRAGMENT).
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
[1]
RN RP SEQUENCE FROM N.A.
RA Ryder T.B., Cramer C.L., Bell J.N., Robbins M.P., Dixon R.A.,
RA Lamb C.J.;
RT "Elicitor rapidly induces chalcone synthase mRNA in Phaseolus vulgaris
RT cells at the onset of the phytoalexin defense response.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5724-5728 (1984).
DR EMBL; K02953; AAA33758.1; -.
DR InterPro; IPR001099; Chal_stil_synth.
DR Pfam; PF00195; Chal_stil_synth; 1.
FT NON_TER 1
FT 75
SQ SEQUENCE 75 AA; 8800 MW; 243CBED889A4F207 CRC64;

```

```

Query Match 91.7%; Score 22; DB 10; Length 75;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   || ||
Db 18 GTATPS 23

```

```

RESULT 3
Q90QH8 ID Q90QH8 PRELIMINARY; PRT; 80 AA.
AC Q90QH8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN RP SEQUENCE FROM N.A.
RA Kaufmann G.R., Suzuki K., Cunningham P., Mukaide M., Kondo M.,
RA Inai M., Zaunders J., Cooper D.A.;
RT "Impact of HIV Type 1 Protease, Reverse Transcriptase, Cleavage Site,
RT and p6 Mutations on the Virological Response to Quadraple Therapy with
RT Saquinavir, Ritonavir, and Two Nucleoside Analogs.";
RL AIDS Res. Hum. Retroviruses 17:487-497 (2001).
DR EMBL; AF322213; AAK66674.1; -.
FT NON_TER 1
FT 80
SQ SEQUENCE 80 AA; 8942 MW; 72C5C3148E51E65D CRC64;

```

```

Query Match 91.7%; Score 22; DB 15; Length 80;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   || ||
Db 48 GTTTPS 53

```

```

RESULT 4
Q9HWZ4 ID Q9HWZ4 PRELIMINARY; PRT; 89 AA.
AC Q9HWZ4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA4033.
GN PA4033.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004820; AAG07420.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 89 AA; 9452 MW; BFD4EBF98897E7E4 CRC64;

```

```

Query Match 91.7%; Score 22; DB 16; Length 89;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
   || ||
Db 48 GTASPS 53

```

```

RESULT 5
O38158 ID O38158 PRELIMINARY; PRT; 92 AA.
AC O38158;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=PATIENT 4;
RC MEDLINE=97404676; PubMed=9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
RT the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670 (1997).
DR EMBL; AF024054; AAB83833.1; -.
DR HSP; P05888; IAAF.
DR InterPro; IPR001878; znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
DR SMART; SM00343; znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT 92
SQ SEQUENCE 92 AA; 10403 MW; 27E1160C29F611EA CRC64;

```

```

Query Match 91.7%; Score 22; DB 15; Length 92;

```


Best Local Similarity 66.7%; Pred. No. 1.3e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 4; Conservative 0;

QY 1 gtxxps 6
|| ||
Db 72 GTTPS 77

RESULT 6

O38175 PRELIMINARY; PRT; 93 AA.
AC O38175;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;
RX MEDLINE=97404676; PubMed=9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
RT the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024071; AAB83850.1; -.
DR HSSP; P05888; IAAF.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; Znf_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 2.
KW zinc-finger.
FT NON_TER 1 93
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10402 MW; 0E9FF41DC4129C73 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 93;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 73 GTATPS 78

RESULT 7

O38177 PRELIMINARY; PRT; 93 AA.
AC O38177;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;
RX MEDLINE=97404676; PubMed=9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
RT the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024074; AAB83852.1; -.
DR HSSP; P05888; IAAF.
DR InterPro; IPR001878; Znf_CCHC.

DR Pfam; PF00098; Znf_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 2.
KW zinc-finger.
FT NON_TER 1 93
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10502 MW; 5615F69550DDA09D CRC64;

Query Match 91.7%; Score 22; DB 15; Length 93;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 73 GTATPS 78

RESULT 8

O38114 PRELIMINARY; PRT; 96 AA.
AC O38114;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;
RX MEDLINE=97404676; PubMed=9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
RT the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024010; AAB83789.1; -.
DR HSSP; P05888; IAAF.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; Znf_CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW zinc-finger.
FT NON_TER 1 96
FT NON_TER 96 96
SQ SEQUENCE 96 AA; 10766 MW; 74A041FFD4FE3935 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 96;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 76 GTATPS 81

RESULT 9

O38161 PRELIMINARY; PRT; 96 AA.
AC O38161;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;

RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024057; AAB83836.1; -.
 DR HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 96 96
 SQ SEQUENCE 96 AA; 10824 MW; 74B74DFE73E93935 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 96;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||
 Db 76 GTTTPS 81

RESULT 10
 O38178 PRELIMINARY; PRT; 96 AA.
 ID O38178
 AC O38178;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024075; AAB83853.1; -.
 DR HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 96 96
 SQ SEQUENCE 96 AA; 10794 MW; 74B756FFC3E93935 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 96;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||
 Db 76 GTTTPS 81

RESULT 11
 O38030 PRELIMINARY; PRT; 99 AA.
 ID O38030
 AC O38030;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 3;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF023926; AAB83860.1; -.
 DR HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 99 99
 SQ SEQUENCE 99 AA; 11140 MW; 92348443928FDCC0 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 99;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||
 Db 79 GTTTPS 84

RESULT 12
 O38362 PRELIMINARY; PRT; 99 AA.
 ID O38362
 AC O38362;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 7;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024261; AAB83135.1; -.
 DR HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 99 99
 SQ SEQUENCE 99 AA; 11063 MW; 9818C494BED11470 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 99;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||
 Db 79 GTTTPS 84

RESULT 12
 O38362 PRELIMINARY; PRT; 99 AA.
 ID O38362
 AC O38362;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 7;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024261; AAB83135.1; -.
 DR HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 99 99
 SQ SEQUENCE 99 AA; 11063 MW; 9818C494BED11470 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 99;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || || ||

```
Db 79 GTATPS 84

RESULT 13
O38371 ID O38371 PRELIMINARY; PRT; 99 AA.
AC O38371;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PATIENT 7; PubMed=9261388;
RC MEDLINE=97404676; Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RAS Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024270; AAB83144.1; -
DR HSSP; P05888; 1AAF.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT NON_TER 99
FT NON_TER 99
SQ SEQUENCE 99 AA; 11092 MW; 9818C494BD1E1470 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 79 GTATPS 84

RESULT 14
O38372 ID O38372 PRELIMINARY; PRT; 99 AA.
AC O38372;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PATIENT 7; PubMed=9261388;
RC MEDLINE=97404676; Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RAS Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024271; AAB83145.1; -
DR HSSP; P05888; 1AAF.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT NON_TER 99
FT NON_TER 99
SQ SEQUENCE 99 AA; 11092 MW; 9818C494BD1E1470 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 79 GTATPS 84

RESULT 15
O38113 ID O38113 PRELIMINARY; PRT; 101 AA.
AC O38113;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PATIENT 4; PubMed=9261388;
RC MEDLINE=97404676; Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RAS Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024009; AAB83788.1; -
DR HSSP; P05888; 1AAF.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT NON_TER 101
FT NON_TER 101
SQ SEQUENCE 101 AA; 11408 MW; 2EF1A3E5CF282854 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 81 GTTTPS 86

RESULT 16
O38128 ID O38128 PRELIMINARY; PRT; 101 AA.
AC O38128;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PATIENT 4; PubMed=9261388;
RC MEDLINE=97404676; Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
RAS Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024271; AAB83145.1; -
DR HSSP; P05888; 1AAF.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT NON_TER 101
FT NON_TER 101
SQ SEQUENCE 101 AA; 11408 MW; 2EF1A3E5CF282854 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
DB 81 GTTTPS 86
```

RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024024; AAB83803.1; -
 DR HSP; P05888; IAAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 101 101
 SQ SEQUENCE 101 AA; 11295 MW; EBD5B3278BF1D5FA CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gtxxps 6
 || ||
 Db 81 GTTTPS 86

RESULT 17
 O38130
 ID O38130 PRELIMINARY; PRT; 101 AA.
 AC O38130;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024026; AAB83805.1; -
 DR HSP; P05888; IAAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 101 101
 SQ SEQUENCE 101 AA; 11408 MW; F7B36D60472BDE0 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gtxxps 6
 || ||
 Db 81 GTTTPS 86

RESULT 18
 O38133
 ID O38133 PRELIMINARY; PRT; 101 AA.
 AC O38133;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024029; AAB83808.1; -
 DR HSP; P05888; IAAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 101 101
 SQ SEQUENCE 101 AA; 11364 MW; F7B9D83A5781D5FA CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gtxxps 6
 || ||
 Db 81 GTATPS 86

RESULT 19
 O38138
 ID O38138 PRELIMINARY; PRT; 101 AA.
 AC O38138;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 RT the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024034; AAB83813.1; -
 DR HSP; P05888; IAAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 101 101
 SQ SEQUENCE 101 AA; 11478 MW; 02BB523A5781CEFB CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gtxxps 6
 || ||
 Db 81 GTATPS 86

RESULT 20
 O38141

```

ID 038141 PRELIMINARY; PRT; 101 AA.
AC 038141;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;
RX MEDLINE-97404676; PubMed-9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
  Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
  the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024037; AAB83816.1; -.
DR HSSP; P05888; 1AAF.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam; PF000098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT NON_TER 101
FT SEQUENCE 101 AA; 11379 MW; F7BB523A5781D5FA CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
  || ||
Db 81 GTATPS 86

RESULT 21
Q9YA39 PRELIMINARY; PRT; 105 AA.
AC Q9YA39;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 11.3 KDA PROTEIN APE2100.
GN APE2100.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-K1;
RX MEDLINE-99310339; PubMed-10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
  Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
  Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
  Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
  Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
  Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
  crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000063; BAA81111.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 105 AA; 11262 MW; B25D94F6CED05D87 CRC64;

Query Match 91.7%; Score 22; DB 17; Length 105;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 gtxxps 6
  || ||
Db 15 GTSSPS 20

RESULT 22
O38165 PRELIMINARY; PRT; 111 AA.
ID 038165;
AC 038165;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;
RX MEDLINE-97404676; PubMed-9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
  Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
  the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024061; AAB83840.1; -.
DR HSSP; P05888; 1AAF.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam; PF000098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.
KW Zinc-finger.
FT NON_TER 1
FT NON_TER 111
FT SEQUENCE 111 AA; 12601 MW; A3611A55B709F506 CRC64;

```

```

Query Match 91.7%; Score 22; DB 15; Length 111;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
  || ||
Db 91 GTTTPS 96

RESULT 23
O38166 PRELIMINARY; PRT; 111 AA.
ID 038166;
AC 038166;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-PATIENT 4;
RX MEDLINE-97404676; PubMed-9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
  Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in
  the protease gene and in its Gag substrate cleavage sites.";
RL J. Virol. 71:6662-6670(1997).
DR EMBL; AF024062; AAB83841.1; -.
DR HSSP; P05888; 1AAF.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam; PF000098; zf-CCHC; 2.
DR SMART; SM00343; Znf_C2HC; 2.

```

KW Zinc-finger.
 FT NON_TER 111 1
 SQ SEQUENCE 111 AA; 12684 MW; FC3BEFE25A1B7EDB CRC64;

Query Match 91.7%; Score 22; DB 15; Length 111;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 91 GTTTPS 96

RESULT 24
 O38169
 ID O38169 PRELIMINARY; PRT; 111 AA.
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024065; AAB83844.1; -
 DR HSP; P05888; 1AAF.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 111 1
 SQ SEQUENCE 111 AA; 12684 MW; 39C82CAA0ADFDE CRC64;

Query Match 91.7%; Score 22; DB 15; Length 111;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 91 GTTTPS 96

RESULT 25
 O38171
 ID O38171 PRELIMINARY; PRT; 111 AA.
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,

RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024067; AAB83846.1; -
 DR HSP; P05888; 1AAF.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 111 1
 SQ SEQUENCE 111 AA; 12482 MW; 653BE64427F9D8FC CRC64;

Query Match 91.7%; Score 22; DB 15; Length 111;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 91 GTTTPS 96

RESULT 26
 Q75698
 ID Q75698 PRELIMINARY; PRT; 112 AA.
 AC Q75698;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GAG POLYPROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96213510; PubMed=8638406;
 RA Barrie K.A., Perez E.E., Lammers S.L., Farmerie W.G., Dunn B.M.,
 Sleasman J.W., Goodenow M.M.;
 RT "Natural variation in HIV-1 protease, Gag p7 and p6, and protease
 cleavage sites within gag/pol polyproteins: amino acid substitutions
 in the absence of protease inhibitors in mothers and children infected
 by human immunodeficiency virus type 1.";
 RL Virology 219:407-416(1996).
 DR EMBL; U53663; AAB05739.1; -
 DR HSP; P05888; 1AAF.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Polyprotein; Zinc-finger.
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12895 MW; C97021F2A4407CA7 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 112;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 || ||
 Db 94 GTTTPS 99

RESULT 27
 Q64053
 ID Q64053 PRELIMINARY; PRT; 135 AA.
 AC Q64053;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE ORF 5' OF FIBROBLAST GROWTH FACTOR RECEPTOR 1 FGFR-1.
GN FGFR1.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95100926; PubMed=7802632;
RA Harada T., Saito H., Kouchi H., Kurebayashi S., Kasayama S.,
RA Terakawa N., Kishimoto T., Sato B.;
RT "Murine fibroblast growth factor receptor 1 gene generates multiple
RT messenger RNAs containing two open reading frames via alternative
RT splicing."
RL Biochem. Biophys. Res. Commun. 205:1057-1063(1994).
DR EMBL: S74765; AAB32844.2; -
DR MGD; MGI:95522; Fgfr1.
KW Receptor.
SQ SEQUENCE 135 AA; 14280 MW; 03906AD6DC40880F CRC64;

Query Match 91.7%; Score 22; DB 11; Length 135;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 105 GTAAPS 110

RESULT 28
Q9D5C6 PRELIMINARY; PRT; 142 AA.
ID Q9D5C6
AC Q9D5C6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 4930465A12RIK PROTEIN.
GN 4930465A12RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK015502; BAB29872.1; -
DR MGD; MGI:192239; 4930465A12RIK.
SQ SEQUENCE 142 AA; 15397 MW; AF4B8868A17B745E CRC64;

Query Match 91.7%; Score 22; DB 11; Length 142;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 74 GTATPS 79

RESULT 29
Q9HBN7 PRELIMINARY; PRT; 156 AA.
ID Q9HBN7
AC Q9HBN7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 16.2 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.X., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
RT growth."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF218021; AAG17263.1; -
KW Hypothetical protein.
SQ SEQUENCE 156 AA; 16178 MW; 183180BEC8F93E63 CRC64;

Query Match 91.7%; Score 22; DB 4; Length 156;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
DB 10 GTAAPS 15

RESULT 30
Q9NWG3 PRELIMINARY; PRT; 157 AA.
ID Q9NWG3
AC Q9NWG3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 17.4 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Tsogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK009095; BAA91418.1; -
SQ SEQUENCE 157 AA; 17352 MW; 2B1C874775BC2D23 CRC64;

Query Match 91.7%; Score 22; DB 4; Length 157;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 76 GTAPs 81

RESULT 32
O41995
ID Q41995 PRELIMINARY; PRT; 177 AA.
AC O41995;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
OS maize ravado fino virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Marafivirus.
OX NCBI_TaxID=59749;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UNITED STATES;
RX MEDLINE=98062136; PubMed=9400964;
RA Hammond R.W., Kogel R., Ramirez P.;
RT "Variability of geographically distinct isolates of maize rayado fino
virus in Latin America";
RL J. Gen. Virol. 78:0-0(0).
DR EMBL: U97729; AAB96574.1;
DR InterPro: IPR000574; Tymo_coat.
DR Pfam: PF00983; Tymo_coat; 1.
FT NON_TER 1
SQ SEQUENCE 177 AA; 18918 MW; F1A95271D011795E CRC64;

Query Match 91.7%; Score 22; DB 12; Length 177;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
DB 171 GTAPs 176

RESULT 33
Q9KY82
ID Q9KY82 PRELIMINARY; PRT; 178 AA.
AC Q9KY82;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 19.7 KDA PROTEIN.
GN SCK15.22.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL356813; CAB92613.1;
KW Hypothetical protein.
SQ SEQUENCE 178 AA; 19681 MW; 155BA91389A2DDE6 CRC64;

Db 76 GTAPs 81

RESULT 31
ID Q9VS07 PRELIMINARY; PRT; 175 AA.
AC Q9VS07;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE CG13309 PROTEIN.
GN CG13309.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram P.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003554; AAF50359.1;
DR FlyBase: FBgn0035933; CG13309.
SQ SEQUENCE 175 AA; 18038 MW; A094BA899ADFALC7 CRC64;

Query Match 91.7%; Score 22; DB 5; Length 175;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gtxxps 6
DB 98 GTAPs 103

Query Match 91.7%; Score 22; DB 2; Length 178;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 2 GTSTPS 7

RESULT 34

Q9VSO6 PRELIMINARY; PRT; 182 AA.
 AC Q9VSO6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG13308 PROTEIN.
 GN CG13308.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yangell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003554; AAF50360.1;
 DR FlyBase; FBgn0035932; CG13308;
 SQ SEQUENCE 182 AA; 18808 MW; 64B9775C44A3AEE2 CRC64;

Query Match 91.7%; Score 22; DB 5; Length 182;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 105 GTSTPS 110

RESULT 35

Q9BZ94 PRELIMINARY; PRT; 186 AA.
 AC Q9BZ94;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ZINC FAMILY MEMBER 4 PROTEIN HZIC4.
 GN ZIC4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MEDULLOBLASTOMA;
 RA Warder D.E.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF332509; AAK06845.1;
 DR HSSP; P08047; ISP2.
 DR InterPro; IPR000822; Znf-C2H2.
 DR Pfam; PF00096; zf-C2H2; 1.
 DR SMART; SM00355; Znf-C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 2.
 KW DNA-binding; Metal-binding; Zinc-finger.
 SQ SEQUENCE 186 AA; 20189 MW; BF4C37753D290527 CRC64;

Query Match 91.7%; Score 22; DB 4; Length 186;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 67 GTATPS 72

RESULT 36

O82664 PRELIMINARY; PRT; 196 AA.
 AC O82664;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BLUE COPPER BINDING-LIKE PROTEIN (AT5G20230/F5024_120) (PUTATIVE BLUE
 COPPER BINDING PROTEIN).
 GN AMI 32 OR ATBCD.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang K.Y., Kim C.S., Cho B.H.;
 RT "Characterization of a wound-inducible Arabidopsis gene encoding a
 RT protein homologous to blue copper binding proteins";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. LANSBERG ERECTA;
 RA Homma T., Goto K.;
 RT "The Arabidopsis floral homeotic gene PISTILLATA is regulated by
 RT discrete cis-elements responsive to induction and maintenance
 RT signals";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yanamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cdna clones";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F5024.120/AT5g20230.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y18227; CAA77089.1; -;
 DR EMBL; AB035137; BAA86999.1; -;
 DR EMBL; AY052681; AK96585.1; -;
 DR EMBL; AY034986; AAK59491.1; -;
 DR HSP; P29602; IJER.
 DR InterPro; IPR000923; Copper_blue1.
 DR InterPro; IPR003245; Cu_bind_like.
 DR Pfam; PF02298; Cu_bind_like; 1.
 DR ProDom; PD003122; Cu_bind_like; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 SQ SEQUENCE 196 AA; 20053 MW; 05100B50518F0A56 CRC64;

Query Match 91.7%; Score 22; DB 10; Length 196;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 158 GTTTPS 163

RESULT 37
 ID Q9F5I3 PRELIMINARY; PRT; 200 AA.
 AC Q9F5I3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE RIOP4 PROTEIN.
 GN RIOP4.
 OS Agrobacterium rhizogenes.
 OG Plasmid pRi1724.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RA Moriguchi K., Maeda Y., Satou M., Satuti N., Kataoka M., Tanaka N.,
 RA Yoshida K.;
 RT "The complete nucleotide sequence of a Ri (root inducing) plasmid
 RT indicates its chimerical structure between Ti and Sym plasmids.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RA Moriguchi K., Maeda Y., Satou M., Kataoka M., Tanaka N., Yoshida K.;
 RA "Analysis of unique variable region of a plant root inducing plasmid,
 RT pRi1724, by the construction of its physical map and library.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RA Moriguchi K., Nishida T., Maeda Y., Tanaka N., Yoshida K.;
 RT "Genome structure of Ri plasmid (1).Construction of linking library
 RT and physical map of pRi1724 in Japanese Agrobacterium.";
 RL Nucleic Acids Symp. Ser. 39:189-190(1998).
 [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MAFF03-01724;
 RC MEDLINE=20241294; PubMed=10780382;
 RA Maeda Y., Moriguchi K., Kataoka M., Satou M., Satuti N., Tanaka N.,
 RA Yoshida K.;
 RT "Genome structure of Ri plasmid (1): Sequencing analysis of T-DNA and
 RT its flanking regions of pRi1724 in Japanese Agrobacterium
 RT rhizogenes.";
 RL Nucleic Acids Symp. Ser. 42:67-68(1999).
 CC -1- SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.
 DR EMBL; AP002086; BAB16123.1; -;
 DR InterPro; IPR000847; HTH_Lysr.
 DR Pfam; PF00126; HTH_1; 1.
 KW DNA-binding; Plasmid; Transcription regulation.
 SQ SEQUENCE 200 AA; 21360 MW; 68888FA7699BFCEAF CRC64;

Query Match 91.7%; Score 22; DB 2; Length 200;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 177 GTAAPS 182

RESULT 38
 ID Q9KOL1 PRELIMINARY; PRT; 202 AA.
 AC Q9KOL1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE OUTER MEMBRANE LIPOPROTEIN SLP, PUTATIVE.
 GN VC1987.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004273; AAF95135.1; -;
 DR TIGR; VC1987; -;
 KW Lipoprotein; Complete proteome.
 SQ SEQUENCE 202 AA; 22661 MW; DA43B8A8D40DB691 CRC64;

Query Match 91.7%; Score 22; DB 16; Length 202;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||

Db 134 GTTAPS 139

RESULT 39

Q9JLH5 PRELIMINARY; PRT; 218 AA.
 AC Q9JLH5
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE CKK5 ACTIVATOR-BINDING PROTEIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20184747; PubMed=10721722;
 RA Ching Y.P., Qi Z., Wang J.H.;
 RT "Cloning of three novel neuronal Cdk5 activator binding proteins.";
 RL Gene 242:285-294(2000).
 DR EMBL; AF177478; AAF60224.1; -;
 SQ SEQUENCE 218 AA; 24023 MW; 2C6750AABA0D0EDA CRC64;

Query Match 91.7%; Score 22; DB 11; Length 218;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 209 GTSSPS 214

RESULT 40

Q9RDK2 PRELIMINARY; PRT; 228 AA.
 AC Q9RDK2
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE PUTATIVE MEMBRANE PROTEIN.
 GN SC123.22C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RA Brown S.P., Harris D.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL136518; CAB66262.1; -;
 SQ SEQUENCE 228 AA; 24246 MW; 6217C85671050B30 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 228;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 209 GTSSPS 214

Qy 1 gtxxps 6
 || ||
 Db 206 GTATPS 211

RESULT 41

Q9KXV0 PRELIMINARY; PRT; 233 AA.
 AC Q9KXV0
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE PUTATIVE SECRETED PROTEIN.
 GN SCD95A.24.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL357432; CAB93051.1; -;
 SQ SEQUENCE 233 AA; 23071 MW; D3E0BEE9A356083D CRC64;

Query Match 91.7%; Score 22; DB 2; Length 233;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 52 GTATPS 57

RESULT 42

P79289 PRELIMINARY; PRT; 237 AA.
 AC P79289
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE SP1 TRANSCRIPTION FACTOR (FRAGMENT).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-ENDOMETRIUM;
 RL MEDLINE=20150569; PubMed=10687861;
 RA Simmen R.C.M., Zhang X.L., Zhang D., Wang Y., Michel F.J.,
 RA Simmen F.A.;
 RT "Expression and regulatory function of the transcription factor Sp1 in
 the uterine endometrium at early pregnancy: Implications for
 epithelial phenotype.";
 RL Mol. Cell. Endocrinol. 159:159-170(2000).

DR EMBL; U57347; AAB39513.3; -.
 DR HSSP; P08047; ISP2.
 DR InterPro; IPR000822; Znf-C2H2.
 DR Pfam; PF00096; ZF-C2H2_3.
 DR SMART; SM00355; ZNF-C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 KW DNA-binding; Metal-binding; Zinc-finger.
 FT NON_TER 1
 FT NON_TER 237
 SQ SEQUENCE 237 AA; 25421 MW; C6950DB42912DAB6 CRC64;

Query Match 91.7%; Score 22; DB 6; Length 237;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 204 GTATPS 209

RESULT 43
 Q9L203 PRELIMINARY; PRT; 247 AA.
 AC Q9L203;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PUTATIVE BIOTIN SYNTHASE.
 GN SC8E4.05C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL138661; CAB71805.1; -.
 GN INTERPRO: IPR003784; BioY.
 DR Pfam; PF02632; BioY; 1.
 SQ SEQUENCE 247 AA; 24640 MW; 3A5EF99FDD7CF0CE CRC64;

Query Match 91.7%; Score 22; DB 2; Length 247;
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 150 GTAAPS 155

RESULT 44
 Q9L126 PRELIMINARY; PRT; 253 AA.
 ID Q9L126
 AC Q9L126;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE PUTATIVE LIPOPROTEIN.
 GN SC6D11.30.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL158061; CAB76353.1; -.
 KW Lipoprotein.
 SQ SEQUENCE 253 AA; 26600 MW; FECC3325A465D891 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 253;
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 || ||
 Db 122 GTASPS 127

RESULT 45
 O43563 PRELIMINARY; PRT; 253 AA.
 ID O43563
 AC O43563;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE ORCL2S HYPOTHETICAL PROTEIN.
 GN ORCL2S OR BWRIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cooper P.R., Smilnich N.J., Day C.D., Nowak N.J., Reid L.H.,
 RA Pearsall R.S., Reece M., Prawitt D., Landers J., Housman D.E.,
 RA Winterpacht A., Zabel B.U., Pelletier J., Weissman B.E., Shows T.B.,
 RA Higgins M.J.;
 RL Genomics 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98188297; PubMed=9520460;
 RA Schwenbacher C., Sabbioni S., Campi M., Veronese A., Bernardi G.,
 RA Menegatti A., Hatada I., Mukai T., Ohashi H., Barbanti-Brodano G.,
 RA Croce C.M., Negrini M.;
 RT "Transcriptional map of 170-kb region at chromosome 11p15.5:
 identification and mutational analysis of the BWR1A gene reveals the
 presence of mutations in tumor samples.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3873-3878(1998).
 DR EMBL; AF037066; AAC04789.1; -.

DR EMBL: AF035407; AAC17497.1; -
SQ SEQUENCE 253 AA; 27218 MW; 1183DE7B5C8A2255B CRC64;

Query Match 91.7%; Score 22; DB 4; Length 253;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 240 GTSTPS 245

RESULT 46

Q91170 ID Q91170 PRELIMINARY; PRT; 254 AA.
AC Q91170;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE PROBABLE THIOESTERASE.
GN PA2411.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.,
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
DR EMBL: AE004668; AAG05799.1; -
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00975; Thioesterase; 1.
KW Complete proteome.
SQ SEQUENCE 254 AA; 27878 MW; 460C521723BD7403 CRC64;

Query Match 91.7%; Score 22; DB 16; Length 254;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 108 GTAAPS 113

RESULT 47

Q98TY7 ID Q98TY7 PRELIMINARY; PRT; 275 AA.
AC Q98TY7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOD.
GN MYOD.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Gao J., Du S.;
RT "Isolation and Characterization of Muscle-Specific Expression of

RT Zebrafish Myod Gene."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH AN OTHER
CC BHLH PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL: AF318503; AAK06755.1; -
DR HSSP; P10085; IMDY.
DR InterPro: IPR002546; Basic.
DR InterPro: IPR001092; HLH_dim.
DR InterPro: IPR003015; HLH_myc.
DR Pfam: PF01586; Basic; 1.
DR Pfam: PF00010; HLH; 1.
DR SMART; SM00520; BASIC; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW DNA-binding.
SQ SEQUENCE 275 AA; 30918 MW; 0A395542F95B37D8 CRC64;

Query Match 91.7%; Score 22; DB 13; Length 275;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 253 GTTAPS 258

RESULT 48

Q9M0A9 ID Q9M0A9 PRELIMINARY; PRT; 285 AA.
AC Q9M0A9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 30.9 KDA PROTEIN.
GN AT4G30510.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161577; CAB79769.1; -
DR InterPro: IPR001680; WD40.
DR SMART; SM00320; WD40; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 285 AA; 30935 MW; E8E7017F494969E1 CRC64;

Query Match 91.7%; Score 22; DB 10; Length 285;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
|| ||
Db 233 GTSSPS 238

RESULT 49

Q9RNC9 ID Q9RNC9 PRELIMINARY; PRT; 288 AA.
AC Q9RNC9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE VIRB9 HOMOLOG.
GN VIRB9.
OS Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=38323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOUSTON-1;
RX MEDLINE=20338084; PubMed=10882236;
RA Padmalayam I., Karem K., Baumstark B., Massung R.;
RT "The gene encoding the 17-kDa antigen of Bartonella henselae is
RT located within a cluster of genes homologous to the virB virulence
RT operon."
RL DNA Cell Biol. 19:377-382(2000).
DR EMBL: AF182718; AAF00947.1; -.
SQ SEQUENCE 288 AA; 32120 MW; B94F15C58F38D489 CRC64;

Query Match          91.7%; Score 22; DB 2; Length 288;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 271 GTTSPS 276

RESULT 50
Q9DWD6
ID Q9DWD6 PRELIMINARY; PRT; 288 AA.
AC Q9DWD6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE PF58.
GN R58.
OS Rat cytomegalovirus (strain Maastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=79700;
RN [1]
RP SEQUENCE OF 1-168 FROM N.A.
RC STRAIN=MAASTRICHT;
RX MEDLINE=96335691; PubMed=8757999;
RA Beuken E., Slobbe R., Bruggeman C.A., Vink C.;
RT "Cloning and sequence analysis of the genes encoding DNA polymerase,
RT glycoprotein B, ICP18.5 and major DNA-binding protein of rat
RT cytomegalovirus."
RL J. Gen. Virol. 77:1559-1562(1996).
RN [2]
RP SEQUENCE OF 167-288 FROM N.A.
RC STRAIN=MAASTRICHT;
RX MEDLINE=98033207; PubMed=9367384;
RA Vink C., Beuken E., Bruggeman C.A.;
RT "Cloning and functional characterization of the origin of lytic-phase
RT DNA replication of rat cytomegalovirus."
RL J. Gen. Virol. 78:2963-2973(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAASTRICHT;
RX MEDLINE=20366325; PubMed=10906222;
RA Vink C., Beuken E., Bruggeman C.A.;
RT "Complete DNA sequence of the rat cytomegalovirus genome."
RL J. Virol. 74:7656-7665(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAASTRICHT;
RX MEDLINE=20473137; PubMed=11018281;
RA Gruijthuisen Y.K., Beuken E., Bruggeman C.A., Vink C.;
RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
```

```
RT spliced transcript."
RL Virus Res. 69:119-130(2000).
DR EMBL: AF232689; AAF99154.1; -.
SQ SEQUENCE 288 AA; 30131 MW; C8F592D7837C9A34 CRC64;

Query Match          91.7%; Score 22; DB 12; Length 288;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 247 GTATPS 252

Search completed: September 24, 2002, 11:26:15
Job time: 724 sec
```

OM of: BASK-853-CLAIM5 to: GenEmbl.* out_format : pfs

Date: Sep 24, 2002 12:09 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+2n.model -DEV=xlp
-Q/cn2_1/uspto.spool/BASK853079/runat_17092002_103407_1819/app_query.fasta_1.98
-DB=GenEmbl -OFF=fastap -SUFFIX=mod.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-CGAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=50 -DOALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=BASK853079.@CGL1.1.6129 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: BASK-853-CLAIM5

Query length: 6

Database: GenEmbl.*

Database sequences: 1797656

Database length: 187333701

Search time (sec): 1991.950000

Sequence list:

Sequence	Strid Orig	ZScore	EScore	Len	Documentation
gb_pat:AX128307	+	22.00	111.17	34	AX128307 Sequence 74 from Patent
gb_pat:E16240	+	22.00	111.17	34	E16240 PCR primer for mouse GT g
gb_pat:E16285	+	22.00	111.17	34	E16285 PCR primer for mouse GT g
gb_ro:AX155976	+	22.00	108.45	51	AX155976 Sequence 219 from Patent
gb_ro:AF205720	+	22.00	108.45	51	AF205720 Mus musculus isolate 1
gb_ro:AF328690	+	22.00	108.45	51	AF328690 Mus musculus isolate 1
gb_ro:MMU8389	+	22.00	108.45	51	048389 Mus musculus membrane bou
gb_ro:AF255988	+	22.00	105.20	83	AF255988 Mus musculus clone C7-g
gb_ro:AB01937502	-	22.00	105.12	84	AB019375 Mus musculus gene for m
gb_pr:HSRNP018	-	22.00	104.51	92	AF164650 Homo sapiens tyrosine k
gb_sts:G33091	+	22.00	104.08	98	G33091 Eb84g7S Human chromosome
gb_cm:AF345500	+	22.00	103.75	123	AF345500 Sus scrofa beta-1 adre
gb_pr:AF088756	+	22.00	102.51	124	AF088756 Homo sapiens clone TCR
gb_pr:AF088773	+	22.00	102.51	124	AF088773 Homo sapiens clone TCR
gb_pr:AF088454	+	22.00	102.25	129	AF088454 Homo sapiens clone 164
gb_pl:OSA320264	+	22.00	101.99	134	AF320264 Oryza sativa 2113 snob
gb_pr:HUMADPT22	+	22.00	101.94	135	M29785 Human NAD+ ADP-ribosylth
gb_pr:HS172B8F	-	22.00	101.42	146	257345 H.sapiens CpG island DNA
gb_pr:HUMKER105	+	22.00	101.37	147	M21754 Human pot. pseudo-kerat
gb_ro:MUSIGHAAZ	+	22.00	101.28	149	M33399 Mouse Ig rearranged H-ch
gb_pl:AIN413962	+	22.00	101.24	150	AJ413962 Agropyron intermedium
gb_pr:HS4325043	+	22.00	101.19	151	AJ325043 Homo sapiens genomic s
gb_pr:HS179B38	+	22.00	100.95	159	Z59881 H.sapiens CpG island DNA
gb_pl:AB017746	-	22.00	100.76	161	AB017746 Phomopsis sp. P-JP-24c
gb_pl:AB017742	-	22.00	100.68	163	AB017741 Phomopsis sp. P-A-19
gb_pl:AB017743	-	22.00	100.68	163	AB017742 Fusicoccum sp. No74-1
gb_pl:AB017744	-	22.00	100.68	163	AB017743 Phomopsis sp. 980811-1
gb_pl:AB017745	-	22.00	100.68	163	AB017744 Phomopsis sp. No.1-1
gb_ro:547233	+	22.00	100.40	170	547233 flg-fibroblast growth fa
gb_sts:AU0494330	+	22.00	100.36	171	AX094330 Sequence 51 from patent
gb_sts:AU048602	+	22.00	100.36	171	AU048602 Rattus norvegicus, Ory
gb_pr:AF148053	+	22.00	95.98	181	AF148053 Homo sapiens clone 30r
gb_pr:GCRG21	+	22.00	95.91	183	X00107 Galago Alu repeat type I
gb_pl:WU303115	-	22.00	95.73	188	AJ303115 Hordeum vulgare partia
gb_pr:AX193248	-	22.00	95.21	203	AX193248 Sequence 815 from paten
gb_pr:HUMPRP22	+	22.00	95.21	203	K02578 Human salivary prolinae
gb_pr:HS94D6F	+	22.00	98.95	211	264000 H.sapiens CpG island DNA
gb_pat:AX336622	+	22.00	98.92	212	AX336622 Sequence 7131 from Pat
gb_sts:AF343867	-	22.00	98.83	215	AF343867 Oryza sativa subsp. in

X1 = Gly
X2 = Thr
X5 = Pro

gb_pl:AF006804	-	22.00	98.80	3.9e+03	216	AF006804 Bryonia cretica clo
gb_vi:AF282969 <td>+</td> <td>22.00 <th>98.77</th> <th>3.9e+03</th> <th>217</th> <th>AF282969 HIV-1 isolate 1079</th> </td>	+	22.00 <th>98.77</th> <th>3.9e+03</th> <th>217</th> <th>AF282969 HIV-1 isolate 1079</th>	98.77	3.9e+03	217	AF282969 HIV-1 isolate 1079
gb_pl:AF006805 <td>+</td> <td>22.00 <th>98.73</th> <th>3.9e+03</th> <th>218</th> <th>AF006805 Bryonia cretica clo</th> </td>	+	22.00 <th>98.73</th> <th>3.9e+03</th> <th>218</th> <th>AF006805 Bryonia cretica clo</th>	98.73	3.9e+03	218	AF006805 Bryonia cretica clo
gb_sts:KLJAJ9915 <td>+</td> <td>22.00 <th>98.61</th> <th>3.9e+03</th> <th>222</th> <th>AJ229915 Kluyveromyces lacti</th> </td>	+	22.00 <th>98.61</th> <th>3.9e+03</th> <th>222</th> <th>AJ229915 Kluyveromyces lacti</th>	98.61	3.9e+03	222	AJ229915 Kluyveromyces lacti
gb_pl:AY022702 <td>+</td> <td>22.00 <th>98.55</th> <th>4.0e+03</th> <th>224</th> <th>AY022702 Oryza sativa micros</th> </td>	+	22.00 <th>98.55</th> <th>4.0e+03</th> <th>224</th> <th>AY022702 Oryza sativa micros</th>	98.55	4.0e+03	224	AY022702 Oryza sativa micros
gb_pl:AY022777 <td>+</td> <td>22.00 <th>98.55</th> <th>4.0e+03</th> <th>224</th> <th>AY022777 Oryza sativa micros</th> </td>	+	22.00 <th>98.55</th> <th>4.0e+03</th> <th>224</th> <th>AY022777 Oryza sativa micros</th>	98.55	4.0e+03	224	AY022777 Oryza sativa micros
gb_pl:AY023473 <td>-</td> <td>22.00 <th>98.55</th> <th>4.0e+03</th> <th>224</th> <th>AY023473 Oryza sativa micros</th> </td>	-	22.00 <th>98.55</th> <th>4.0e+03</th> <th>224</th> <th>AY023473 Oryza sativa micros</th>	98.55	4.0e+03	224	AY023473 Oryza sativa micros
gb_ov:AF005431 <td>+</td> <td>22.00 <th>98.49</th> <th>4.0e+03</th> <th>226</th> <th>AF005431 Varanus komodoensis</th> </td>	+	22.00 <th>98.49</th> <th>4.0e+03</th> <th>226</th> <th>AF005431 Varanus komodoensis</th>	98.49	4.0e+03	226	AF005431 Varanus komodoensis
gb_pl:SHVCHS <td>+</td> <td>22.00 <th>98.49</th> <th>4.0e+03</th> <th>226</th> <th>K02953 Phaseolus vulgaris ch</th> </td>	+	22.00 <th>98.49</th> <th>4.0e+03</th> <th>226</th> <th>K02953 Phaseolus vulgaris ch</th>	98.49	4.0e+03	226	K02953 Phaseolus vulgaris ch
gb_pr:1549E3F <td>-</td> <td>22.00 <th>98.38</th> <th>4.1e+03</th> <th>230</th> <th>265624 H.sapiens CpG island</th> </td>	-	22.00 <th>98.38</th> <th>4.1e+03</th> <th>230</th> <th>265624 H.sapiens CpG island</th>	98.38	4.1e+03	230	265624 H.sapiens CpG island

seq_name: gb_pat:AX128307

seq_documentation_block:

LOCUS	AX128307	34 bp	DNA	linear	PAT 15-MAY-2001
DEFINITION	Sequence 74 from Patent WO0130992.				
ACCESSION	AX128307				
VERSION	AX128307.1	GI:14134828			
KEYWORDS	synthetic construct.				
SOURCE	synthetic construct				
ORGANISM	artificial sequence.				
REFERENCE	1 (bases 1 to 34)				
AUTHORS	Koike,C.				
TITLE	g(a)1-3 galactosyltransferase gene and promoter				
JOURNAL	Patent: WO 0130992-A 74 03-MAY-2001.				
	UNIV. PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION				

FEATURES

source	Location/Qualifiers
1..34	
/organism="synthetic construct"	
/db_xref="taxon:32630"	
/note="primer for identifying murine exons 5 and 6"	
6 a	13 c 7 g 8 t

BASE COUNT

ORIGIN

alignment_scores:

Quality	22.00	Length:	6
Ratio:	3.667	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

BASK-853-CLAIM5 x AX128307 ..

Align seg 1/1 to: AX128307 from: 1 to: 34

1 GlyThr*****ProSer 6

|||||

9 GGAACACACAGTCTCT 26

seq_name: gb_pat:E16240

seq_documentation_block:

LOCUS	E16240	34 bp	DNA	linear	PAT 28-JUL-1999
DEFINITION	PCR primer for mouse GT gene.				
ACCESSION	E16240				
VERSION	E16240.1	GI:5710923			
KEYWORDS	JP 1998146193-A/4.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 34)				
AUTHORS	Koike,C.				
TITLE	DNNA CONSTRUCT FOR HOMOLOGOUS RECOMBINATION, SELECTION OF				
JOURNAL	TRANSFORMANT AND HOMOLOGOUS RECOMBINATION USING THE CONSTRUCT				
COMMENT	Patent: JP 1998146193-A 4 02-JUN-1998;				
	KOIKE CHIHIRO				
	OS None				
	OC Artificial sequences.				
	PN JP 1998146193-A/4				
	PD 02-JUN-1998				
	PF 19-APR-1997 JP 1997116372				

```

PR 23-SEP-1996 JP 96P 287242
PI KOIKE CHIHRO
PC C12N15/09, C07H21/04, C07K14/705, C07K16/28, C07K16/34, C12N5/10,
PC C12N9/10,
PC C12Q1/68, G01N33/53, G01N33/53, (C12N15/09, C12R1.91); CC
strandedness: Single;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1..34 /organism='Artificial sequences'
FEATURES
    source
        Location/Qualifiers
            1..34
                /organism='unidentified'
                /db_xref='taxon:32644'
BASE COUNT 7 a 12 c 7 g 8 t
ORIGIN

alignment_scores:
    Quality: 22.00 Length: 6
    Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x E16240 ..
Align seg 1/1 to: E16240 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACACCACTGCTCTCT 26

seq_name: gb_pat:E16240

seq_documentation_block:
LOCUS E16285 34 bp DNA linear PAT 28-JUL-1999
DEFINITION PCR primer for mouse GT gene.
ACCESSION E16285
VERSION E16285.1 GI:5710968
KEYWORDS JP 1998150990-A/4.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 34)
AUTHORS Koike,C.
TITLE NEW GENE TRANSFER
JOURNAL Patent: JP 1998150990-A 4 09-JUN-1998;
COMMENT KOIKE CHIHRO
OS None
OC Artificial sequences.
PN JP 1998150990-A/4
PD 09-JUN-1998
PF 18-NOV-1996 JP 1996322352
PI KOIKE CHIHRO
PC C12N15/09, A01K67/027, C07H21/04, C07K14/42, C12N5/10; CC
strandedness: Single;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1..34 /organism='Artificial sequences'
FEATURES
    source
        Location/Qualifiers
            1..34
                /organism='unidentified'
                /db_xref='taxon:32644'
BASE COUNT 7 a 12 c 7 g 8 t
ORIGIN

alignment_scores:
    Quality: 22.00 Length: 6
    Ratio: 3.667 Gaps: 0

```

```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x E16285 ..
Align seg 1/1 to: E16285 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACACCACTGCTCTCT 26

seq_name: gb_pat:AX155976

seq_documentation_block:
LOCUS AX155976 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 219 from Patent WO0140474.
ACCESSION AX155976
VERSION AX155976.1 GI:14537066
KEYWORDS
SOURCE Chlamydia sp.
ORGANISM Chlamydia sp.
REFERENCE 1 (bases 1 to 51)
AUTHORS Probst,P., Bhatia,A., Skeiky,Y.A., Fling,S.P. and Scholler,J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0140474-A 219 07-JUN-2001;
FEATURES
    source
        Location/Qualifiers
            1..51
                /organism='Chlamydia sp.'
                /db_xref='taxon:35827'
BASE COUNT 16 a 16 c 8 g 11 t
ORIGIN

alignment_scores:
    Quality: 22.00 Length: 6
    Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AX155976 ..
Align seg 1/1 to: AX155976 from: 1 to: 51
1 GlyThr*****ProSer 6
|||||
5 GGTACCGCATCACCATCA 22

seq_name: gb_ro:AF205720

seq_documentation_block:
LOCUS AF205720 51 bp DNA linear ROD 26-JUL-2000
DEFINITION Mus musculus isolate 1.2-66a T-cell receptor beta chain VDJ
ACCESSION AF205720
VERSION AF205720.1 GI:9453938
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 51)
AUTHORS Attuili,V., Bucher,P., Rossi,M., Mutin,M. and Maryanski,J.L.
TITLE Comparative T cell receptor repertoire selection by antigen after adoptive transfer: A glimpse at an antigen-specific preimmune repertoire
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (15), 8473-8478 (2000)
MEDLINE 20359731
REFERENCE 2 (bases 1 to 51)
AUTHORS Attuili,V., Bucher,P., Rossi,M., Mutin,M. and Maryanski,J.L.
TITLE Direct Submission

```

JOURNAL Submitted (17-NOV-1999) INSERM U503 Ecole Normale Supérieur de
Lyon, 46 Allée d'Italie, Lyon, Cedex 07 69364, France

FEATURES

source
Location/Qualifiers
1..51
/organism="Mus musculus"
/strain="DBA/2"
/isolate="1.2-66a"
/db_xref="taxon:10090"
/rearranged
mRNA
<1..>51
/product="T-cell receptor beta chain VDJ junctional
region"
CDS
<1..>51
/note="TCRBV10BJLS2"
/codon_start=1
/product="T-cell receptor beta chain VDJ junctional
region"
/protein_id="AAF87603.1"
/db_xref="GI:9453939"
/translation="SAVYLCASSVGDDYTFG"
BASE COUNT 7 a 16 c 15 g 13 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AF205720 ..

Align seg 1/1 to: AF205720 from: 1 to: 51

1 GlyThr*****ProSer 6
|||||
32 GGGACGACTACACCTCG 49

seq_name: gb_ro:AF328690

seq_documentation_block:

LOCUS AF328690 51 bp DNA linear ROD 29-JUN-2001
DEFINITION Mus musculus isolate 1.2-70b T-cell receptor beta chain VDJ
junctional region gene, partial cds.

ACCESSION AF328690

VERSION AF328690.1 GI:13898542

KEYWORDS

house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Maryanski, J.L., Attali, V., Hamrouni, A., Mutin, M., Rossi, M.,
Aublin, A., and Bucher, P.

TITLE Individuality of Ag-selected and preimmune TCR repertoires

JOURNAL Immunol. Res. 23 (1), 75-84 (2001)

MEDLINE

PUBMED 21310440

REFERENCE

2 (bases 1 to 51)
Maryanski, J.L., Attali, V., Hamrouni, A., Mutin, M., Rossi, M.,
Aublin, A., and Bucher, P.

TITLE

Direct Submission

JOURNAL Submitted (14-DEC-2000) INSERM Unit 503, CERV1, 21 Avenue Tony

Garnier, 69365 Lyon Cedex 07, France

FEATURES

source
Location/Qualifiers
1..51
/organism="Mus musculus"
/strain="DBA/2"
/isolate="1.2-70b"
/db_xref="taxon:10090"
/rearranged
mRNA
<1..>51
/product="T-cell receptor beta chain VDJ junctional

CDS
<1..>51
/note="TCRBV10"
/codon_start=1
/product="T-cell receptor beta chain VDJ junctional
region"
/protein_id="AAK48790.1"
/db_xref="GI:13898543"
/translation="SAVYLCASSYGNDYTFG"
BASE COUNT 9 a 16 c 13 g 13 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AF328690 ..

Align seg 1/1 to: AF328690 from: 1 to: 51

1 GlyThr*****ProSer 6
|||||
32 GGACGACTACACCTCG 49

seq_name: gb_ro:MMU48389

seq_documentation_block:

LOCUS MMU48389 51 bp mRNA linear ROD 05-JUL-1996
DEFINITION Mus musculus membrane bound dipeptidase type I mRNA, 5'

untranslated region.

ACCESSION U48389

VERSION U48389.1 GI:1403697

KEYWORDS

house mouse strain-FVB.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Habib, G.M., Barrios, R., Shi, Z.-Z. and Lieberman, M.W.

TITLE Four distinct membrane-bound dipeptidase RNAs are differentially

expressed and show discordant regulation with gamma-glutamyl

transpeptidase

JOURNAL J. Biol. Chem. 271 (27), 16273-16280 (1996)

MEDLINE

REFERENCE 96279176

2 (bases 1 to 51)

Habib, G.M., Barrios, R., Shi, Z.-Z. and Lieberman, M.W.

AUTHORS

Direct Submission

TITLE

Submitted (05-FEB-1996) Geetha M. Habib, Pathology, Baylor College

of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL

Location/Qualifiers

FEATURES

source
1..51
/organism="Mus musculus"
/strain="FVB"
gene
1..51
/db_xref="taxon:10090"
5'UTR
1..51
/gene="membrane bound dipeptidase type I"
BASE COUNT 9 a 14 c 10 g 17 t 1 others
ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x MMU48389 ..

```

Align seg 1/1 to: MMU48389 from: 1 to: 51
1 GlyThr*****ProSer 6
|||||
1 GGCACAGCAGCTCCATCT 18

seq_name: gb_ro:AF265988

seq_documentation_block:
LOCUS AF265988 83 bp DNA linear ROD 11-OCT-2000
DEFINITION Mus musculus clone C7-97 immunoglobulin heavy chain gene, VDJ
rearrangement region.
ACCESSION AF265988
VERSION AF265988.1 GI:10442307
KEYWORDS
SOURCE
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 83)
Heavy chain revision in MRL mice: A potential mechanism for the
development of autoreactive B cell precursors
J. Immunol. 165 (8), 4487-4493 (2000)
20491940
2 (bases 1 to 83)
Monestier, M. and Klonowski, K.D.
Direct Submission
Submitted (10-MAY-2000) Department of Microbiology and Immunology,
Temple University School of Medicine, 3400 N. Broad St., Rm. 507,
Kresge, Philadelphia, PA 19140, USA
FEATURES
source
1..83
Location/Qualifiers
/organism="Mus musculus"
/strain="C3H/HeJ"
/db_xref="taxon:10090"
/clone="C7-97"
/cell_type="pre-B cell"
/rearranged
<1..>83
/genes="immunoglobulin heavy chain"
misc_feature 1..83
/genes="immunoglobulin heavy chain"
/genes="immunoglobulin heavy chain"
/notes="VDJ rearrangement region"
BASE COUNT 19 a 19 c 26 g 19 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AF265988 ..
Align seg 1/1 to: AF265988 from: 1 to: 83
1 GlyThr*****ProSer 6
|||||
55 GGCACACAGCTCACCCTCT 72

seq_name: gb_ro:AB01937S02

seq_documentation_block:
LOCUS AB01937S02 84 bp DNA linear ROD 14-APR-2000
DEFINITION Mus musculus gene for muCdc7, exon 3.
ACCESSION AB019379
VERSION AB019379.1 GI:3845586
KEYWORDS muCdc7.
SEGMENT 2 of 11
SOURCE Mus musculus liver DNA.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 83)
Heavy chain revision in MRL mice: A potential mechanism for the
development of autoreactive B cell precursors
J. Immunol. 165 (8), 4487-4493 (2000)
20491940
2 (bases 1 to 83)
Monestier, M. and Klonowski, K.D.
Direct Submission
Submitted (10-MAY-2000) Department of Microbiology and Immunology,
Temple University School of Medicine, 3400 N. Broad St., Rm. 507,
Kresge, Philadelphia, PA 19140, USA
FEATURES
source
1..83
Location/Qualifiers
/organism="Mus musculus"
/strain="C3H/HeJ"
/db_xref="taxon:10090"
/clone="C7-97"
/cell_type="pre-B cell"
/rearranged
<1..>83
/genes="immunoglobulin heavy chain"
misc_feature 1..83
/genes="immunoglobulin heavy chain"
/genes="immunoglobulin heavy chain"
/notes="VDJ rearrangement region"
BASE COUNT 19 a 19 c 26 g 19 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AB01937S02/rev ..
Align seg 1/1 to reverse of: AB01937S02 from: 1 to: 84
1 GlyThr*****ProSer 6
|||||
40 GGTACAGCTTCCACCAAGC 23

seq_name: gb_pr:HSRONP018

seq_documentation_block:
LOCUS HSRONP018 92 bp DNA linear PRI 12-JAN-2001
DEFINITION Homo sapiens tyrosine kinase (RON) gene, intron 16.
ACCESSION AF164650
VERSION AF164650.1 GI:9621839
KEYWORDS
SEGMENT 18 of 22
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 92)
Angelson, B.E., Danilkovitch-Milagkova, A., Ivanov, S.V., Breathnach, R.,
Johnson, B.E., Leonard, E.J. and Lerman, M.I.
Gene structure of the human receptor tyrosine kinase RON and
mutation analysis in lung cancer samples
Genes Chromosomes Cancer 29 (2), 147-156 (2000)
20416072
PUBMED 10959094
REFERENCE 2 (bases 1 to 92)
AUTHORS Angeloni, D.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1999) Laboratory of Immunobiology, National
Cancer Institute - Frederick Cancer Research Facility, Bldg. 560 Rm

```

FEATURES
source 12 26, Frederick, MD 21702, USA
Location/Qualifiers
1. .92

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p21.3"

Intron
1. .92
/gene="RON"
/number=16
16 a 29 c 28 g 19 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x HSRONP018/rev ..

Align seg 1/1 to reverse of: HSRONP018 from: 1 to: 92

1 GlyThr*****ProSer 6
|||||
63 GGCACAGCAGCTCTCT 46

seq_name: gb_sts:G33091

seq_documentation_block:

LOCUS G33091 98 bp DNA linear STS 28-SEP-1998
DEFINITION Eb84g7S Human chromosome 11q13 STSS Homo sapiens STS genomic,
Sequence tagged site.

ACCESSION G33091
VERSION G33091.1 GI:2222836
KEYWORDS STS.
SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Guru,S.C., Olufemi,S.-E., Manickam,P., Cummings,C., Gieser,L.M.,
Pike,B.L., Bittner,M.L., Jiang,Y., Chinault,A.C., Nowak,N.J.,
Brzozowska,A., Crabtree,J.S., Wang,Y.P., Roe,B.A., Weismann,J.M.,
Boguski,M.S., Agarwal,S.K., Burns,A.L., Spiegel,A.M., Marx,S.J.,
Flejtner,W.L., de Jong,P.J., Collins,F.S. and Chandrasekharappa,S.C.
A 2.8-Mb clone contig of the multiple endocrine neoplasia type 1
(MEN1) region at 11q13
Genomics 42 (3), 436-445 (1997)

JOURNAL
MEDLINE 9734911
PubMed 9205115

COMMENT

Contact: Siradanahalli C. Guru
Laboratory of Gene Transfer
National Human Genome Research Institute
NIH, Bethesda, MD 20892, USA
Email: sguru@nhgri.nih.gov

Primer A: TCACCTTAACGGCGCGGGACTACATC
Primer B: CCCACCGCGCGCACCTC
STS size: 98

PCR profile:

Presoak: 94 deg C-5min
Denaturation: 94 deg C-30sec
Annealing: 60 deg C-30sec
Polymerization: 72 deg C-1min
PCR cycles: 35
Final extension: 72 deg C-10min.
Location/Qualifiers

FEATURES
source

1. .98
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="11q13"

/clone_lib="Human chromosome 11q13 STSS"
1. .98
primer_bind 12 a 37 c 34 g 15 t
primer_bind complement(80. .98)
BASE COUNT
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x G33091 ..

Align seg 1/1 to: G33091 from: 1 to: 98

1 GlyThr*****ProSer 6
|||||
18 GGGACTACATCTCCAGC 35

seq_name: gb_om:AF345500

seq_documentation_block:

LOCUS AF345500 103 bp DNA linear MAM 24-OCT-2001
DEFINITION Sus scrofa beta-1 adrenergic receptor gene, partial cds.

ACCESSION AF345500
VERSION AF345500.1 GI:14269015
KEYWORDS .
SOURCE pig.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE
AUTHORS Lee,J.H., Zhang,W. and Moran,C.

TITLE Comparative porcine gene mapping relative to human chromosomes 9,
10, 20 and 22

JOURNAL Anim. Genet. 32 (5), 313-315 (2001)

MEDLINE 21540585
PubMed 11683720

REFERENCE 2 (bases 1 to 103)

AUTHORS Lee,J.H., Zhang,W., Lyons,L.A. and Moran,C.

TITLE Direct Submission

JOURNAL Submitted (05-FEB-2001) Animal Science, University of Sydney,
Sydney, NSW 2006, Australia

FEATURES
source Location/Qualifiers

1. .103

/organism="Sus scrofa"

/db_xref="taxon:9823"

<1. .>103

/product="beta-1 adrenergic receptor"

<1. .>103

/note="ADRB1"

/codon_start=3

/product="beta-1 adrenergic receptor"

/protein_id="AAK57992.1"

/db_xref="GI:14269016"

/translation="PFGATIVVGHWDYGSFFCELMTSDVLCVTAS"

BASE COUNT 14 a 30 c 35 g 24 t

ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AF345500 ..

Align seg 1/1 to: AF345500 from: 1 to: 103

```
1 GlyThr*****ProSer 6
|||||
37 GGGACTACGGCTAGCTCTTC 54

seq_name: gb_pr:AF088756

seq_documentation_block:
LOCUS AF088756 124 bp mRNA linear PRI 19-DEC-1998
DEFINITION Homo sapiens clone TCRBV20S1.63 T-cell receptor beta chain
(TCRBV20S1) mRNA, partial cds.
ACCESSION AF088756 GI:4038203
VERSION AF088756.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 124)
AUTHORS Manfras,B.J.
TITLE Rearranged human TCRBV20S1 genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 124)
AUTHORS Manfras,B.J.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1998) Department of Internal Medicine, University
Hospital Ulm, Robert-Koch-Str. 8, Ulm D-89081, Germany
FEATURES
source
1..124
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCRBV20S1.63"
/cell_type="T-lymphocyte"
<1..>124
/gene="TCRBV20S1"
<1..>124
/gene="TCRBV20S1"
/codon_start=1
/product="T-cell receptor beta chain"
/protein_id="AAC97306.1"
/db_xref="GI:4038204"
/translation="LSDSGFYLCARREGLAVPDTQYFGPGTRLTVLEDLKNVFP"

BASE COUNT 27 a 36 c 34 g 27 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AF088756/rev ..
Align seg 1/1 to reverse of: AF088756 from: 1 to: 124

1 GlyThr*****ProSer 6
|||||
53 GGTACCGCTAGTCCTTCC 36

seq_name: gb_pr:AF088756/rev ..

seq_documentation_block:
LOCUS AF088756 124 bp mRNA linear PRI 19-DEC-1998
DEFINITION Homo sapiens clone TCRBV20S1.80 T-cell receptor beta chain
(TCRBV20S1) mRNA, partial cds.
ACCESSION AF088756 GI:4038236
VERSION AF088756.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 124)
```

```
AUTHORS Manfras,B.J.
TITLE Rearranged human TCRBV20S1 genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 124)
AUTHORS Manfras,B.J.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1998) Department of Internal Medicine, University
Hospital Ulm, Robert-Koch-Str. 8, Ulm D-89081, Germany
FEATURES
source
1..124
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCRBV20S1.80"
/cell_type="T-lymphocyte"
<1..>124
/gene="TCRBV20S1"
<1..>124
/gene="TCRBV20S1"
/codon_start=1
/product="T-cell receptor beta chain"
/protein_id="AAC97322.1"
/db_xref="GI:4038237"
/translation="LSDSGFYLCARREGLAVPDTQYFGPGTRLTVLEDLKNVFP"

BASE COUNT 27 a 36 c 34 g 27 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AF088773/rev ..
Align seg 1/1 to reverse of: AF088773 from: 1 to: 124

1 GlyThr*****ProSer 6
|||||
53 GGTACCGCTAGTCCTTCC 36

seq_name: gb_pr:AF088454

seq_documentation_block:
LOCUS AF088454 129 bp mRNA linear PRI 11-MAY-1999
DEFINITION Homo sapiens clone 164 T-cell receptor beta chain (TCRBV10S1P)
mRNA, partial cds.
ACCESSION AF088454 GI:4768546
VERSION AF088454.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 129)
AUTHORS Manfras,B.J.
TITLE Rearranged human TCRBV10S1P genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 129)
AUTHORS Manfras,B.J.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1998) Department of Internal Medicine, University
Hospital Ulm, Robert-Koch-Str. 8, Ulm D-89081, Germany
FEATURES
source
1..129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="164"
/cell_type="T-cell"
<1..>129
/gene="TCRBV10S1P"
<1..>129
/gene="TCRBV10S1P"
```

```

/codon_start=2
/product="T-cell receptor beta chain"
/protein_id="A029548.1"
/db_xref="GI:4768547"
/translation="PVHGVGHSTVFLCQQSQGERLHLRFGDVNRCRGPQGVPT"
BASE COUNT      31 a   35 c   38 g   25 t
ORIGIN

```

```

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM5 x AF088454

```

```

Align seg 1/1 to: AF088454 from: 1 to: 129

```

```

1 GlyThr*****ProSer 6
|||||
61 GGAACGGCTACACCTCG 78

```

```

seq_name: gb_pl:OSA320264

```

```

seq_documentation_block:
LOCUS      OSA320264                      134 bp   mRNA   linear   PLN 10-OCT-2001
DEFINITION Oryza sativa z113 snRNA gene.
ACCESSION  AJ320264
VERSION    AJ320264.1 GI:16075347
KEYWORDS   Z113 small nucleolar RNA; Z113 snRNA gene.
SOURCE     Oryza sativa.
ORGANISM   Oryza sativa.

```

```

REFERENCE   1 (bases 1 to 134)
AUTHORS     Liang, D., Zhang, P., Chen, X., Zhou, H. and Qu, L. H.
TITLE       Identification of Oryza sativa snRNA gene cluster VI
JOURNAL     Unpublished

```

```

REFERENCE   2 (bases 1 to 134)
AUTHORS     Qu, L. H.
TITLE       Direct Submission
JOURNAL     Submitted (23-JUL-2001) Qu L. H., Biotechnology Research Center,
School of Life Science, Zhongshan University, Guangzhou, 510275,
CHINA

```

```

FEATURES             Location/Qualifiers
     source           1..134
                     /organism="Oryza sativa"
                     /db_xref="taxon:4530"
                     /chromosome="2"

```

```

snRNA
     gene            1..134
                     /gene="Z113 snRNA"
                     /product="Z113 small nucleolar RNA"
BASE COUNT      28 a   35 c   35 g   36 t
ORIGIN

```

```

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM5 x OSA320264

```

```

Align seg 1/1 to: OSA320264 from: 1 to: 134

```

```

1 GlyThr*****ProSer 6
|||||
104 GGGACTGCATCTCCATCG 121

```

```

seq_name: gb_pr:HUMADPRT22

```

```

seq_documentation_block:
LOCUS      HUMADPRT22                      135 bp   DNA   linear   PRI 30-OCT-1994
DEFINITION Human NAD+ ADP-ribosyltransferase (ADPRT) gene, exon 22.
ACCESSION  M29785 M22953
VERSION    M29785.1 GI:178187
KEYWORDS   ADP-D-ribosyltransferase; NAD+ ADP-ribosyltransferase.
SEGMENT    22 of 23
SOURCE     Human HeLa cell line, cDNA to mRNA, clones Hz[19,20,21,22,23,24].
and liver DNA.

```

```

ORGANISM   Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE   1 (bases 1 to 125)
AUTHORS     Herzog, H.
JOURNAL     Unpublished (1989)

```

```

REFERENCE   2 (bases 1 to 20; 116 to 135)
AUTHORS     Auer, B., Nagl, U., Herzog, H., Schneider, R. and Schweiger, M.
JOURNAL     Human nuclear NAD+ ADP-ribosyltransferase (polymorizing):
organization of the gene

```

```

TITLE       DNA 8 (8), 575-580 (1989)
JOURNAL     90091744
MEDLINE
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
by H.Herzog, 07-MAR-1989.

```

```

FEATURES             Location/Qualifiers
     source           1..135
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="lq41-q42"

```

```

     intron          <1..10
                     /gene="PPOL"

```

```

     exon            11..125
                     /note="NAD+ ADP-ribosyltransferase, (EC 2.4.2.30; 5' end
put.) G00-119-508; putative"

```

```

     intron          126..>135
                     /number=22
                     /gene="PPOL"
BASE COUNT      29 a   31 c   29 g   46 t
ORIGIN          About 0.5 kb after segment 21; chromosome lq41-q42.

```

```

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM5 x HUMADPRT22/rev

```

```

Align seg 1/1 to reverse of: HUMADPRT22 from: 1 to: 135
1 GlyThr*****ProSer 6
|||||
71 GGAACGCTACACCATCC 54

```

```

seq_name: gb_pr:HS172B8F

```

```

seq_documentation_block:
LOCUS      HS172B8F                      146 bp   DNA   linear   PRI 18-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 172B8,
forward read cpq172b8.ftla.

```

```

ACCESSION  257345
VERSION    257345.1 GI:1028576
KEYWORDS   CpG island; genomic MseI fragment.
SOURCE     human.

```

```

ORGANISM   Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
REFERENCE
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
source
1..146
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="male"
/dev_stage="adult"
/tissue_type="blood"
/clone_lib="CGI-1"
/clone="172b8"
BASE COUNT 27 a 45 c 47 g 24 t 3 others
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x HSL72B8F/rev ..

Align seg 1/1 to reverse of: HSL72B8F from: 1 to: 146

1 GlyThr*****ProSer 6
|||||
110 GGGACTGCCAGCCCTTC 93

seq_name: gb_pr:HUMKERI05

seq_documentation_block:
LOCUS HUMKERI05 147 bp DNA linear PRI 06-JAN-1995
DEFINITION Human pot. pseudo-keratin K16 type I, exon 5.
ACCESSION M21754 M20336
VERSION M21754.1 GI:186721
KEYWORDS pseudogene; type I keratin.
SEGMENT 5 of 8
SOURCE Human epithelial DNA, clone GK-6.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 147)
Rosenberg,M., Raychaudhury,A., Shows,T.B., Le Beau,M.M. and
Fuchs,E.

TITLE A group of type I keratin genes on human chromosome 17:
characterization and expression
JOURNAL Mol. Cell. Biol. 8 (2), 722-736 (1988)
MEDLINE 88174714

FEATURES
source
1..147
/organism="Homo sapiens"
/db_xref="taxon:9606"
order(M21753.1:174..175,1..19)
/number=4
20..145
/gene="K16 pseudogene"
/number=5
35 a 38 c 45 g 29 t
Chromosome 17; 88 bp after segment 4.

BASE COUNT
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x HUMKERI05/rev ..

Align seg 1/1 to reverse of: HUMKERI05 from: 1 to: 147

1 GlyThr*****ProSer 6
|||||
134 GGGACTGCCAGCTCATCT 117

seq_name: gb_ro:MUSIGHAAZ

seq_documentation_block:
LOCUS MUSIGHAAZ 149 bp mRNA linear ROD 27-APR-1993
DEFINITION Mouse Ig rearranged H-chain mRNA V-D-JH4-region, Id+ hybridoma
C16-11d11, partial cds.
ACCESSION M33399
VERSION M33399.1 GI:194634
KEYWORDS D-region; J-region; V-region; immunoglobulin heavy chain; processed
gene.
SOURCE Mouse (strain Balb.b) adult Id+ hybridoma C16-11d11, cDNA to mRNA.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 149)
AUTHORS Borriero,L., Giorgetti,C.A., Smith,G., Landry,D., Selsing,E.,
Zhukovsky,E. and Press,J.L.

TITLE Neonatal and adult primary B cells use the same germ-line V-H and
V-kappa genes in their (T,G)-A-L-specific repertoire
JOURNAL J. Immunol. 144, 583-592 (1990)
MEDLINE 90111072

FEATURES
source
1..149
/organism="Mus musculus"
/db_xref="taxon:10090"
<1..>149
/note="Ig H-chain V-D-JH4-region"
/codon_start=1
/protein_id="AAA37979.1"
/db_xref="GI:553951"
/translation="ITADTSSNTAYLQLSSLTSEDYVYCARAGKGGYAMDYWGOGT
SAPSP"

BASE COUNT 39 a 45 c 35 g 30 t
ORIGIN Chromosome 12.

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x MUSIGHAAZ ..

Align seg 1/1 to: MUSIGHAAZ from: 1 to: 149

1 GlyThr*****ProSer 6
|||||
127 GGAACCTCAGCACCGTCT 144

seq_name: gb_pl:AIN413962

seq_documentation_block:
LOCUS AIN413962 150 bp DNA linear PLN 02-OCT-2001
DEFINITION Agropyron intermedium partial PT gene for inorganic phosphate
transporter, clone 7.
ACCESSION AJ413962


```

VERSION      AJ413962.1  GI:15912383
KEYWORDS     inorganic phosphate transporter; PT gene.
SOURCE       Thinopyrum intermedium.
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Poideae; Triticeae; Thinopyrum.
REFERENCE    1 (bases 1 to 150)
AUTHORS      Davies,T.G.E., Ying,J., Xu,Q., Li,Z. and Gordon-Weeks,R.
TITLE        Analysis of high-affinity phosphate transporter expression in alien
              translocation lines of Chinese winter wheats
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 150)
AUTHORS      Ying,J.
TITLE        Direct Submission
JOURNAL      Submitted (18-SEP-2001) Ying J., Agriculture and Environment
              Division, IACR-Rothamsted, Harpenden, Herts, AL5 2JQ, UNITED
              KINGDOM
FEATURES     Location/Qualifiers
              1..150
                /organism="Thinopyrum intermedium"
                /db_xref="taxon:85679"
                /clone="7"
                /country="China"
                1..150
                /gene="PT"
                <1..>150
                /gene="PT"
                /codon_start=1
                /product="inorganic phosphate transporter"
                /protein_id="CAC88703.1"
                /db_xref="GI:15912384"
                /translation="VFLIDVGRFAIQLGFFFMFLGLAVPHVHTTGNHVGFFV
              VMFSLT"
              18 a 50 c 43 g 39 t
BASE COUNT   18 a 50 c 43 g 39 t
ORIGIN
...
alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM5 x AIN413962/rev
  Align seg 1/1 to reverse of: AIN413962 from: 1 to: 150
  1 GlyThr*****ProSer 6
  |||||
  89 GGCACGGGAGCCCCAGC 72
seq_name: gb_pr: HSA325043
seq_documentation_block:
LOCUS      HSA325043
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
              NLI-EF8R.
ACCESSION  AJ325043
VERSION    AJ325043
KEYWORDS   human.
SOURCE     Homo sapiens
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 151)
AUTHORS    Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F.,
              Podowski,R.W., Macushkin,I.G., Kvasha,S.M., Gyanchandani,A.,
              Muravenko,O.V., Protopopov,A.I., Kashuba,V.I., Kisselev,L.L.,
              Wassenman,W., Wahlestedt,C. and Zabarovsky,E.R.
TITLE      Analysis of NotI flanking sequences: a new tool for gene discovery
              and verification of the human genome
JOURNAL    Unpublished

```

```

REFERENCE    2 (bases 1 to 151)
AUTHORS      Zabarovsky,E.R.
TITLE        Direct Submission
JOURNAL      Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
              Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
              Sweden
FEATURES     Location/Qualifiers
              1..151
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="NLI-EF8R"
              21 a 50 c 65 g 15 t
BASE COUNT   21 a 50 c 65 g 15 t
ORIGIN
...
alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
  BASK-853-CLAIM5 x HSA325043
  Align seg 1/1 to: HSA325043 from: 1 to: 151
  1 GlyThr*****ProSer 6
  |||||
  117 GGCACGGGAGCCCCAGC 134
seq_name: gb_pr: HSL79B3R
seq_documentation_block:
LOCUS      HSL79B3R
DEFINITION H.sapiens CpG island DNA genomic Msel fragment, clone 179b3,
              reverse read cpgl79b3.r1la.
ACCESSION  Z59881
VERSION    Z59881.1 GI:1031794
KEYWORDS   CpG island; genomic Msel fragment.
SOURCE     human.
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 159)
AUTHORS    Dodsworth,S.J., Huckle,E., Wilkinson,P. and Mickle,M.G.
TITLE      Direct Submission
JOURNAL    Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
REFERENCE  2 (bases 1 to 159)
AUTHORS    Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE      Purification of CpG islands using a methylated DNA binding column
              Nat. Genet. 6 (3), 236-244 (1994)
              94282070
              Vector: pGEM-5zf(-)
              Clones are available from the UK MRC Human Genome Mapping Project
              Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
              http://www.hgmp.mrc.ac.uk/ for details
              or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES     Location/Qualifiers
              1..159
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /sex="male"
                /dev_stage="adult"
                /tissue_type="blood"
                /clone_lib="CGI-1"
                /clone="179b3"
              26 a 55 c 60 g 18 t
BASE COUNT   26 a 55 c 60 g 18 t
ORIGIN
...
alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0

```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x HS179B3R ..

Align seg 1/1 to: HS179B3R from: 1 to: 159

1 GlyThr*****ProSer 6

|||||

121 GGTACAGCGCGCCAGC 138

seq_name: gb_pl:AB017746

seq_documentation_block:

LOCUS AB017746

DEFINITION Phomopsis sp. P-JP-24d DNA, ITS2 region. linear PLN 07-FEB-2001

ACCESSION AB017746

VERSION AB017746.1 GI:4519395

KEYWORDS

SOURCE

ORGANISM Phomopsis sp. P-JP-24d (strain:P-JP-24d) DNA.

REFERENCE 1 (sites) Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Diaporthales; Valsaceae; mitosporic Valsaceae; Phomopsis.

AUTHORS Kanematsu,S.

TITLE rDNA ITS region sequence of Phomopsis

JOURNAL Published Only in DataBase (1999) in press

REFERENCE 2 (bases 1 to 161)

AUTHORS Kanematsu,S.

TITLE Direct Submission

JOURNAL Submitted (18-SEP-1998) Satoko Kanematsu, National Institute of Fruit Tree Science, Dept. of Plant Protection, Plant Pathology; Fujimoto 2-1, Tsukuba, Ibaraki 305-8605, Japan

(E-mail:satokok@affrc.go.jp, Tel:81-298-38-6546)

FEATURES

source

1..161

/organism="Phomopsis sp. P-JP-24d"

/strain="P-JP-24d"

/db_xref="taxon:151284"

misc_feature 1..161

/note="ITS2"

BASE COUNT 37 a 49 c 42 g 33 t

ORIGIN

alignment_scores:

Quality: 22.00

Ratio: 3.667

Length: 6

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AB017746/rev ..

Align seg 1/1 to reverse of: AB017746 from: 1 to: 161

1 GlyThr*****ProSer 6

|||||

132 GGCACCGCCAGTCCTCC 115

seq_name: gb_pl:AB017741

seq_documentation_block:

LOCUS AB017741

DEFINITION Phomopsis sp. P-A-19 DNA, ITS2 region. linear PLN 06-FEB-2001

ACCESSION AB017741

VERSION AB017741.1 GI:4519390

KEYWORDS

SOURCE

ORGANISM Phomopsis sp. P-A-19 (strain:P-A-19) DNA.

REFERENCE 1 (sites) Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Diaporthales; Valsaceae; mitosporic Valsaceae; Phomopsis.

AUTHORS Kanematsu,S.

TITLE rDNA ITS region sequence of Phomopsis

JOURNAL Published Only in DataBase (1999) In press

REFERENCE 2 (bases 1 to 163)

AUTHORS Kanematsu,S.

TITLE Direct Submission

JOURNAL Submitted (18-SEP-1998) Satoko Kanematsu, National Institute of

Fruit Tree Science, Dept. of Plant Protection, Plant Pathology;

Fujimoto 2-1, Tsukuba, Ibaraki 305-8605, Japan

(E-mail:satokok@affrc.go.jp, Tel:81-298-38-6546)

FEATURES

source

1..163

/organism="Phomopsis sp. P-A-19"

/strain="P-A-19"

/db_xref="taxon:151281"

misc_feature 1..163

/note="ITS2"

BASE COUNT 38 a 47 c 45 g 33 t

ORIGIN

alignment_scores:

Quality: 22.00

Ratio: 3.667

Length: 6

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AB017741/rev ..

Align seg 1/1 to reverse of: AB017741 from: 1 to: 163

1 GlyThr*****ProSer 6

|||||

134 GGCACCGCCAGTCCTCC 117

seq_name: gb_pl:AB017742

seq_documentation_block:

LOCUS AB017742

DEFINITION Fusicoccum sp. No74-1 DNA, ITS2 region. linear PLN 03-NOV-1999

ACCESSION AB017742

VERSION AB017742.1 GI:4519391

KEYWORDS

SOURCE

ORGANISM Fusicoccum sp. No74-1 (strain:No. 74-1) DNA.

Fusicoccum sp. No74-1

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et Chaetothyriomycetes incertae sedis; Botryosphaeriaceae; mitosporic Botryosphaeriaceae; Fusicoccum.

REFERENCE 1 (bases 1 to 163)

AUTHORS Kanematsu,S.

TITLE rDNA ITS region sequence of Phomopsis

JOURNAL Published Only in DataBase (1999) In press

REFERENCE 2 (bases 1 to 163)

AUTHORS Kanematsu,S.

TITLE Direct Submission

JOURNAL Submitted (18-SEP-1998) to the DDBJ/EMBL/GenBank databases. Satoko

Kanematsu, National Institute of Fruit Tree Science, Dept. of Plant

Protection, Plant Pathology; Fujimoto 2-1, Tsukuba, Ibaraki

305-8605, Japan (E-mail:satokok@affrc.go.jp, Tel:81-298-38-6546)

FEATURES

source

1..163

/organism="Fusicoccum sp. No74-1"

/strain="No. 74-1"

/db_xref="taxon:83178"

misc_feature 1..163

/note="ITS2"

BASE COUNT 38 a 47 c 45 g 33 t

ORIGIN

alignment_scores:

Quality: 22.00

Ratio: 3.667

Length: 6

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:
BASK-853-CLAIM5 x AB017742/rev ..
Align seg 1/1 to reverse of: AB017742 from: 1 to: 163
1 GlyThr*****ProSer 6
|||||
134 GGCACCGCCAGTCTTCC 117

seq_name: gb_pl:AB017743
seq_documentation_block:
LOCUS AB017743 163 bp DNA linear PLN 06-FEB-2001
DEFINITION Phomopsis sp. 980811-14 DNA, ITS2 region.
ACCESSION AB017743
VERSION AB017743.1 GI:4519392
KEYWORDS
SOURCE
ORGANISM
Phomopsis sp. 980811-14 (strain:980811-14) DNA.
Phomopsis sp. 980811-14;
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Diaporthales; Valsaceae; mitosporic Valsaceae; Phomopsis.
REFERENCE
AUTHORS Kanematsu,S.
TITLE rDNA ITS region sequence of Phomopsis
JOURNAL
REFERENCE 2 (bases 1 to 163)
AUTHORS Kanematsu,S.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1998) Satoko Kanematsu, National Institute of
Fruit Tree Science, Dept. of Plant Protection, Plant Pathology;
Fujimoto 2-1, Tsukuba, Ibaraki 305-8605, Japan
(E-mail:satokok@affrc.go.jp, Tel:81-298-38-6546)
FEATURES
source
1..163
/organism="Phomopsis sp. No.1-1"
/strain="No. 1-1"
/db_xref="taxon:151279"
misc_feature 1..163
BASE COUNT 38 a 47 c 45 g 33 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AB017744/rev ..
Align seg 1/1 to reverse of: AB017744 from: 1 to: 163
1 GlyThr*****ProSer 6
|||||
134 GGCACCGCCAGTCTTCC 117

seq_name: gb_pl:AB017745
seq_documentation_block:
LOCUS AB017745 163 bp DNA linear PLN 03-NOV-1999
DEFINITION Fusicoccum sp. No36-1 DNA, ITS2 region.
ACCESSION AB017745
VERSION AB017745.1 GI:4519394
KEYWORDS
SOURCE
ORGANISM
Fusicoccum sp. No36-1 (strain:No. 36-1) DNA.
Fusicoccum sp. No36-1
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyriomycetes incertae sedis; Botryosphaeriaceae; mitosporic
Botryosphaeriaceae; Fusicoccum.
REFERENCE
AUTHORS Kanematsu,S.
TITLE rDNA ITS region sequence of Phomopsis
JOURNAL
REFERENCE 2 (bases 1 to 163)
AUTHORS Kanematsu,S.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1998) to the DDBJ/EMBL/GenBank databases. Satoko
kanematsu, National Institute of Fruit Tree Science, Dept. of Plant
Protection, Plant Pathology; Fujimoto 2-1, Tsukuba, Ibaraki
305-8605, Japan (E-mail:satokok@affrc.go.jp, Tel:81-298-38-6546)
FEATURES
source
1..163
/organism="Fusicoccum sp. No36-1"
/strain="No. 36-1"
/db_xref="taxon:83177"
misc_feature 1..163
BASE COUNT 38 a 48 c 43 g 34 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
BASK-853-CLAIM5 x AB017745/rev ..
Align seg 1/1 to reverse of: AB017745 from: 1 to: 163

1 GlyThr*****ProSer 6
|||||
134 GGCACCGCCAGTCCTCC 117

seq_name: gb_ro:S47233

seq_documentation_block:
LOCUS S47233 170 bp mRNA linear ROD 08-MAY-1993
DEFINITION flg-fibroblast growth factor receptor FGFR1 [5' region] [mice, MM14
myoblasts, mRNA Partial, 170 nt].
ACCESSION S47233
VERSION S47233.1 GI:258957
KEYWORDS
SOURCE Mus sp. MM14 myoblasts.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 170)
Templeton,T.J. and Hauschka,S.D.
FGF-mediated aspects of skeletal muscle growth and differentiation
are controlled by a high affinity receptor, FGFR1
Dev. Biol. 154 (1), 169-181 (1992)
93050774
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 117204] from the original journal article.
This sequence comes from Fig. 5.
FEATURES
source 1..170
/organism="Mus sp."
/db_xref="taxon:10095"
1..170
/partial
/genes="flg"
/note="fibroblast growth factor receptor FGFR1"
BASE COUNT 40 a 58 c 57 g 15 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x S47233 ..
Align seg 1/1 to: S47233 from: 1 to: 170

1 GlyThr*****ProSer 6
|||||
110 GGCACCGCCAGCCCAAGT 127

seq_name: gb_pat:AR094330

seq_documentation_block:
LOCUS AR094330 171 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 51 from patent US 6001639.
ACCESSION AR094330
VERSION AR094330.1 GI:10021159
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
1 (bases 1 to 171)
Unclassified.
REFERENCE
1 (bases 1 to 171)
Schulein,M., Andersen,L.Nonboe, Lassen,S.slashed.renFlensted,
Kauppinen,M.Sakari, Lange,L., Nielsen,R.Iium, Ihara,M. and
Takagi,S.
Endoglucanases

JOURNAL Patent: US 6001639-A 51 14-DEC-1999;
FEATURES Location/Qualifiers
source 1..171
/organism="unknown"
BASE COUNT 33 a 67 c 50 g 21 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AR094330 ..
Align seg 1/1 to: AR094330 from: 1 to: 171

1 GlyThr*****ProSer 6
|||||
26 GGACGGACAACCCCTCT 43

seq_name: gb_sts:AU048602

seq_documentation_block:
LOCUS AU048602 171 bp DNA linear STS 20-JAN-2000
DEFINITION Rattus norvegicus, OTSUKA clone, 734a09, microsatellite sequence,
sequence tagged site.
ACCESSION AU048602
VERSION AU048602.1 GI:6721773
KEYWORDS STS.
SOURCE Rattus norvegicus (strain:Brown Norway) liver hepatocyte DNA,
clone:734a09.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (sites)
Watanabe,T.K., Hishigaki,H., Okuno,S., Mizoguchi,A., Oga,K.,
Tsuji,A., Ono,T., Yamasaki,Y., Kanemoto,N., Takahashi,E., Irie,Y.,
Nakamura,Y., Takagi,Y. and Tanigami,A.
The large-scale mapping of rat microsatellite markers
Unpublished (1998)
2 (bases 1 to 171)
Watanabe,T.K.
Direct Submission
K Watanabe, Otsuka GEN Research Institute, Otsuka Pharmaceutical
Co., Ltd; 463-10, Kagasuno, Kawauchi-cho, Tokushima, Tokushima
771-0192, Japan (E-mail:watanabe@otsuka.gr.jp, Tel:+81-886-65-2888,
Fax:+81-886-37-1035)
FEATURES Location/Qualifiers
source 1..171
/organism="Rattus norvegicus"
/strain="Brown Norway"
/db_xref="taxon:10116"
/cell_type="hepatocyte"
/clone="734a09"
/tissue_type="liver"
/note="734a09F-5'-CACACATATACAGACACAG-3',
734a09R-5'-GCAGATGGAGTAGAGTACC-3'"
BASE COUNT 63 a 38 c 28 g 42 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AU048602 ..

```

```
Align seg 1/1 to: AU048602 from: 1 to: 171
1 GlyThr*****ProSer 6
101 GGTACTTCTACTCCATCT 118

seq_name: gb_pr:AF148053

seq_documentation_block:
LOCUS AF148053 181 bp DNA linear PRI 12-MAY-2000
DEFINITION Homo sapiens clone 30r t(14;18) breakpoint D1-7/BCL-2 gene fusion
sequence.
ACCESSION AF148053
VERSION AF148053.1 GI:5478660
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 181)
Jager, U., Bocskor, S., Le, T., Mitterbauer, G., Bolz, I., Chott, A.,
Kneba, M., Mannhalter, C. and Nadel, B.
Follicular lymphomas' BCL-2/IgH junctions contain templated
nucleotide insertions: novel insights into the mechanism of
t(14;18) translocation
t(14;18) translocation
Blood 95 (11), 3520-3529 (2000)
JOURNAL 20287393
MEDLINE 2 (bases 1 to 181)
Jaeger, U., Bocskor, S., Le, T., Mitterbauer, G., Bolz, I., Chott, A.,
Kneba, M., Mannhalter, C. and Nadel, B.
Direct Submission
TITLE Submitted (04-MAY-1999) Internal Medicine 1, Division of
Hematology, University of Vienna, Waehringer Guertel 18-20, Vienna
A-1090, Austria
FEATURES
source
Location/Qualifiers
1..181
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="t(14;18)"
/clone="30r"
/notes="t(14;18) reciprocal breakpoint D1-7/BCL-2 gene
fusion"
misc_feature 1..181
BASE COUNT 60 a 36 c 37 g 48 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x GCREG1 ..
Align seg 1/1 to: GCREG1 from: 1 to: 183
1 GlyThr*****ProSer 6
121 GGCACCTCTAGTCCGACG 138

seq_name: gb_pl:HVU303115

seq_documentation_block:
LOCUS HVU303115 188 bp mRNA linear PLN 23-DEC-2000
DEFINITION Hordeum vulgare partial mRNA for lectin, clone 4A-HL#2-R.
ACCESSION AJ303115
VERSION AJ303115.1 GI:11991495
KEYWORDS hl#2 gene; lectin.
barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE
AUTHORS 1 (bases 1 to 188)
Grunwald, I.
Thesis (2001) Department of Biological Sciences, University of
Hannover, Hannover, Germany
REFERENCE
AUTHORS 2 (bases 1 to 188)
Kloppstech, K.R.
Direct Submission
TITLE Submitted (17-OCT-2000) Kloppstech K.R., Institute of Botany,
University of Hannover, Herrenhaeuser Str. 2, 30419 Hannover,
GERMANY
COMMENT under investigation.
FEATURES
source
Location/Qualifiers
1..188
/organism="Hordeum vulgare"
/variety="apex"
/db_xref="taxon:4513"
/clone="4A-HL#2-R"
/tissue_type="etiolated coleoptile"
/dev_stage="day 6"
1..188
/gene="HL#2"
<1..>188
/gene="HL#2"
/codon_start=2
/product="putative lectin"
/protein_id="CAC19672.1"
/db_xref="GI:11991496"
/translation="AVGRDAAGHDPAWRAGRTWKSSVWSDLRLGGSSKRLNSGSR
SRSSASFRGAKMWTSCR"
BASE COUNT 26 a 61 c 79 g 22 t
```

```
Align seg 1/1 to: AU048602 from: 1 to: 171
1 GlyThr*****ProSer 6
101 GGTACTTCTACTCCATCT 118

seq_name: gb_pr:AF148053

seq_documentation_block:
LOCUS AF148053 181 bp DNA linear PRI 12-MAY-2000
DEFINITION Homo sapiens clone 30r t(14;18) breakpoint D1-7/BCL-2 gene fusion
sequence.
ACCESSION AF148053
VERSION AF148053.1 GI:5478660
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 181)
Jager, U., Bocskor, S., Le, T., Mitterbauer, G., Bolz, I., Chott, A.,
Kneba, M., Mannhalter, C. and Nadel, B.
Follicular lymphomas' BCL-2/IgH junctions contain templated
nucleotide insertions: novel insights into the mechanism of
t(14;18) translocation
t(14;18) translocation
Blood 95 (11), 3520-3529 (2000)
JOURNAL 20287393
MEDLINE 2 (bases 1 to 181)
Jaeger, U., Bocskor, S., Le, T., Mitterbauer, G., Bolz, I., Chott, A.,
Kneba, M., Mannhalter, C. and Nadel, B.
Direct Submission
TITLE Submitted (04-MAY-1999) Internal Medicine 1, Division of
Hematology, University of Vienna, Waehringer Guertel 18-20, Vienna
A-1090, Austria
FEATURES
source
Location/Qualifiers
1..181
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="t(14;18)"
/clone="30r"
/notes="t(14;18) reciprocal breakpoint D1-7/BCL-2 gene
fusion"
misc_feature 1..181
BASE COUNT 60 a 36 c 37 g 48 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AF148053 ..
Align seg 1/1 to: AF148053 from: 1 to: 181
1 GlyThr*****ProSer 6
10 GGACTACTCTACCTCTCC 27

seq_name: gb_pr:GCREG1

seq_documentation_block:
LOCUS GCREG1 183 bp DNA linear PRI 20-MAY-1992
DEFINITION Galago Alu repeat type II, GAL 1.
ACCESSION X00107
VERSION X00107.1 GI:22826
KEYWORDS Alu repetitive sequence; repetitive sequence.
thick-tailed bush baby;
SOURCE Otollemur crassicaudatus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Otollemur.
```

ORIGIN

alignment_scores:
 Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 BASK-853-CLAIM5 x HVU303115/rev ..

Align seg 1/1 to reverse of: HVU303115 from: 1 to: 188

1 GlyThr*****ProSer 6
 |||||

182 GGCACGCGTCACCATCT 165

seq_name: gb_pat:AX193248

seq_documentation_block:

LOCUS AX193248 203 bp DNA linear PAT 15-AUG-2001

DEFINITION Sequence 815 from Patent WO0149716.

ACCESSION AX193248

VERSION AX193248.1 GI:15211199

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 203)

Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,

Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.

Compounds for immunotherapy and diagnosis of colon cancer and

methods for their use

Patent: WO 0149716-A 815 12-JUL-2001;

CORIXA CORPORATION (US)

FEATURES

Location/Qualifiers

1..203

/organism="Homo sapiens"

/db_xref="taxon:9606"

46 a 65 c 58 g 34 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AX193248/rev ..

Align seg 1/1 to reverse of: AX193248 from: 1 to: 203

1 GlyThr*****ProSer 6
 |||||

122 GGCACGCGTCACCATCA 105

seq_name: gb_pr:HUMPRP22

seq_documentation_block:

LOCUS HUMPRP22 203 bp DNA linear PRI 27-APR-1993

DEFINITION Human salivary proline-rich protein 2 gene, segment 2.

ACCESSION K02578

VERSION K02578.1 GI:190477

KEYWORDS proline-rich protein.

SEGMENT 2 of 2

SOURCE Human DNA (library of Blattner), clone PRP2.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 203)

Azen,E.A., Lyons,K.M., McGonigal,T., Barrett,N.L., Clements,L.S.,

TITLE

JOURNAL
 MEDLINE
 COMMENT

FEATURES

source

1..203

/organism="Homo sapiens"

/db_xref="taxon:9606"

<1..>140

/codon_start=2

/product="salivary proline-rich protein 2"

/protein_id="AAA36505.1"

/db_xref="GI:553623"

/translation="SARSPRKFGPPQGGNNPPGPPPPAGGNPQQPQAPPAGQPOG

PP"

BASE COUNT 54 a 84 c 42 g 23 t

ORIGIN About 850 bp after segment 1; chromosome 12.

alignment_scores:

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x HUMPRP22 ..

Align seg 1/1 to: HUMPRP22 from: 1 to: 203

1 GlyThr*****ProSer 6

|||||

130 GGCACACCGCCCTCC 147

seq_name: gb_pr:HS94D6F

seq_documentation_block:

LOCUS HS94D6F 211 bp DNA linear PRI 22-OCT-1995

DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 94D6, forward

read cp994d6.ftla.

ACCESSION Z64000

VERSION Z64000.1 GI:1036378

KEYWORDS CpG island; genomic MseI fragment.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 211)

AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Mickle,M.G.

TITLE Direct Submission

JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,

CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk

REFERENCE 2 (bases 1 to 211)

AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.

TITLE Purification of CpG islands using a methylated DNA binding column

JOURNAL Nat. Genet. 6 (3), 236-244 (1994)

MEDLINE 94282070

COMMENT Vector: pGEM-5zf(-)

Clones are available from the UK MRC Human Genome Mapping Project

Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:

http://www.hgmp.mrc.ac.uk/ for details

or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES

Location/Qualifiers

1..211

/organism="Homo sapiens"

/db_xref="taxon:9606"

/sex="male"

/dev_stage="adult"

/tissue_type="blood"

Maeda,N., Vanin,E.F., Carlson,D.M. and Smithies,O.
 Clones from the human gene complex coding for salivary proline-rich
 proteins
 Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565 (1984)
 84298176

The sequence of the PRP2 gene presented here has one 63 bp repeat
 starting at position 11.

Draft entry and sequence on magnetic tape kindly provided by

N.Maeda, June 1985.

Location/Qualifiers

1..203

/organism="Homo sapiens"

/db_xref="taxon:9606"

<1..>140

/product="salivary proline-rich protein 2"

/protein_id="AAA36505.1"

/db_xref="GI:553623"

/translation="SARSPRKFGPPQGGNNPPGPPPPAGGNPQQPQAPPAGQPOG

PP"

BASE COUNT 54 a 84 c 42 g 23 t

ORIGIN About 850 bp after segment 1; chromosome 12.

```

/clone_lib="CGI-1"
/clone="94d6"
BASE COUNT      39 a   73 c   54 g   41 t   4 others
ORIGIN

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x HS94D6F/rev
Align seg 1/1 to reverse of: HS94D6F from: 1 to: 211

1 GlyThr*****ProSer 6
|||||
187 GGGACTAGCTCCCGC 170

seq_name: gb_pat:AX336622

seq_documentation_block:
LOCUS      AX336622      212 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION Sequence 7131 from Patent WO0194629.
ACCESSION  AX336622
VERSION     AX336622.1 GI:18127341
KEYWORDS
SOURCE
  ORGANISM      Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (sites)
AUTHORS      Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
  Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE        Cancer gene determination and therapeutic screening using signature
  gene sets
JOURNAL      Patent: WO 0194629-A 7131 13-DEC-2001;
  Avalon Pharmaceuticals (US)
FEATURES
  source
    1. .212
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT      61 a   36 c   29 g   86 t
ORIGIN

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AX336622
Align seg 1/1 to: AX336622 from: 1 to: 212

1 GlyThr*****ProSer 6
|||||
180 GGCACCTAGCTCTCCCTCA 197

seq_name: gb_sts:AF343867

seq_documentation_block:
LOCUS      AF343867      215 bp      DNA      linear      STS 13-AUG-2001
DEFINITION Oryza sativa subsp. indica microsatellite marker RM82, sequence
  tagged site.
ACCESSION  AF343867
VERSION     AF343867.1 GI:14194031
KEYWORDS
SOURCE
  ORGANISM      Oryza sativa (indica cultivar-group).
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Enarthroideae; Oryzaceae; Oryza.
1 (bases 1 to 215)
AUTHORS      Chen,X., Temnykh,S., Xu,Y., Cho,Y.G. and McCouch,S.R.
TITLE        Development of microsatellite framework map providing genome-wide
  coverage in rice (Oryza sativa L.)
JOURNAL      Theor. Appl. Genet. 95, 553-567 (1997)
REFERENCE
  2 (bases 1 to 215)
AUTHORS      Temnykh,S., DeClerck,G., Lukashova,A., Lipovich,L., Cartinhour,S.
  and McCouch,S.
TITLE        Computational and Experimental Analysis of Microsatellites in Rice
  (Oryza sativa L.): Frequency, Length Variation, Transposon
  Associations, and Genetic Marker Potential
JOURNAL      Genome research. 11 (8), 1441-1452 (2001)
MEDLINE
PUBMED
11483586
REFERENCE
  3 (bases 1 to 215)
AUTHORS      Chen,X., Temnykh,S., Xu,Y., Cho,Y.G. and McCouch,S.R.
TITLE        Direct Submission
JOURNAL      Submitted (27-JAN-2001) Department of Plant Breeding, Cornell
  University, Bradfield Hall, Ithaca, NY 14853-1901, USA
FEATURES
  source
    1. .215
    /organism="Oryza sativa (indica cultivar-group)"
    /cultivar="IR36"
    /sub_species="indica"
    /db_xref="taxon:39946"
    /chromosome="7"
    /clone="TCT122"
    /clone_lib="sheared genomic"
    primer_bind      25. .43
    STS
    primer_bind      25. .139
    /standard_name="microsatellite marker RM82"
BASE COUNT      34 a   74 c   53 g   50 t   4 others
ORIGIN

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AF343867/rev
Align seg 1/1 to reverse of: AF343867 from: 1 to: 215

1 GlyThr*****ProSer 6
|||||
128 GGTACGGCTGCNCGTCN 111

seq_name: gb_pl:AF006804

seq_documentation_block:
LOCUS      AF006804      216 bp      DNA      linear      PLN 15-JUN-1998
DEFINITION Bryonia cretica clone 1 internal transcribed spacer 1, complete
  sequence.
ACCESSION  AF006804
VERSION     AF006804.1 GI:3220071
KEYWORDS
SOURCE
  ORGANISM      Bryonia cretica.
  Bryonia cretica
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
REFERENCE
  1 (bases 1 to 216)
AUTHORS      Jobst,J., King,K. and Hemleben,V.
TITLE        Molecular evolution of the internal transcribed spacers (ITS1 and
  ITS2) and phylogenetic relationships among species of the family
  Cucurbitaceae
JOURNAL      Mol. Phylogenet. Evol. 9 (2), 204-219 (1998)
MEDLINE
98224077

```

```

REFERENCE      2 (bases 1 to 216)
AUTHORS        Jobst,J., King,K. and Hemleben,V.
TITLE          Direct Submission
JOURNAL        Submitted (05-JUN-1997) Genetics, Biology, Auf der Morgenstelle 28,
               Tuebingen 72076, Germany
FEATURES       Location/Qualifiers
                source
                1..216
                /organism="Bryonia cretica"
                /db_xref="taxon:61884"
                /clone="1"
                1..216
                /note="ITS1"
                /product="internal transcribed spacer 1"
BASE COUNT    42 a 77 c 58 g 39 t
ORIGIN
alignment_scores
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AF006804/rev ..
Align seg 1/1 to reverse of: AF006804 from: 1 to: 216
1 GlyThr*****ProSer 6
|||||
81 GGCACGAGTCTCCTCT 64
seq_name: gb_vi:AF282969

seq_documentation_block:
LOCUS      AF282969      217 bp DNA linear VRL 19-SEP-2001
DEFINITION HIV-1 isolate 1079 from Angola gag protein (gag) gene and pol
            protein (pol) gene, partial cds.
ACCESSION  AF282969
VERSION    AF282969.1 GI:9454467
KEYWORDS   Human immunodeficiency virus type 1.
SOURCE     Human immunodeficiency virus type 1
ORGANISM   Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE  1 (bases 1 to 217)
            Peters,S., Munoz,M., Yerly,S., Sanchez-Merino,V.,
            Lopez-Galindez,C., Perrin,L., Larder,B., Cmarko,D., Fakan,S.,
            Meylan,P. and Telenti,A.
            Resistance to Nucleoside Analog Reverse Transcriptase Inhibitors
            Mediated by Human Immunodeficiency Virus Type 1 p6 Protein
            J. Virol. 75 (20), 9644-9653 (2001)
MEDLINE    21443958
PUBMED     11559796
REFERENCE  2 (bases 1 to 217)
            Peters,S., Martinez,R. and Telenti,A.
            Polymorphism of p1/p6 GagPol in HIV-1 subtype B
            Unpublished
JOURNAL    3 (bases 1 to 217)
            Peters,S., Martinez,R. and Telenti,A.
            Direct Submission
TITLE      Submitted (28-JUN-2000) Division of Infectious Diseases, Centre
            Hospitalier Universitaire Vaudois, rue du Bagnon, Lausanne, Vaud
            1011, Switzerland
            Location/Qualifiers
            1..217
            /organism="Human immunodeficiency virus type 1"
            /isolate="1079"
            /db_xref="taxon:11676"
            /country="Angola"
            /note="subtype: B"
            <1..>217
            /gene="pol"
            <1..217

gene
gene

```

```

CDS
/genes="gag"
<1..217
/gene="gag"
/notes="p6; pl"
/codon_start=2
/product="gag protein"
/protein_id="AAF87830.1"
/db_xref="GI:9454469"
/translation="FLGKIWPNSRGRGNFLOSRPEPTAPPAPPEESFRFGECTATPS
OKQEPIDKEMYPLASLSRSLFGNDPSLQ"
<1..>217
/gene="pol"
/codon_start=1
/product="pol protein"
/protein_id="AAF87829.1"
/db_xref="GI:9454468"
/translation="FFREDLAFQOGKAREFSSEQTNTSFTSPTRRKQLQWGRDSNSL
SEAGADRGQGVFSFQITLMQRPVITI"
BASE COUNT    63 a 60 c 51 g 43 t
ORIGIN

alignment_scores
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AF282969 ..
Align seg 1/1 to: AF282969 from: 1 to: 217
1 GlyThr*****ProSer 6
|||||
116 GGCACGACCACTCCCTCT 133
seq_name: gb_pl:AF006805

seq_documentation_block:
LOCUS      AF006805      218 bp DNA linear PLN 15-JUN-1998
DEFINITION Bryonia cretica clone 2 internal transcribed spacer 1, complete
            sequence.
ACCESSION  AF006805
VERSION    AF006805.1 GI:3220072
KEYWORDS   Bryonia cretica.
SOURCE     Bryonia cretica.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
REFERENCE  1 (bases 1 to 218)
            Jobst,J., King,K. and Hemleben,V.
            Molecular evolution of the internal transcribed spacers (ITS1 and
            ITS2) and phylogenetic relationships among species of the family
            Cucurbitaceae
JOURNAL    Mol. Phylogenet. Evol. 9 (2), 204-219 (1998)
MEDLINE    98224077
PUBMED     98224077
REFERENCE  2 (bases 1 to 218)
            Jobst,J., King,K. and Hemleben,V.
            Direct Submission
TITLE      Submitted (05-JUN-1997) Genetics, Biology, Auf der Morgenstelle 28,
            Tuebingen 72076, Germany
            Location/Qualifiers
            1..218
            /organism="Bryonia cretica"
            /db_xref="taxon:61884"
            /clone="2"
            1..218
            /note="ITS1"
            /product="internal transcribed spacer 1"
BASE COUNT    44 a 76 c 59 g 39 t
ORIGIN

```



```

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AF006805/rev

  Align seg 1/1 to reverse of: AF006805 from: 1 to: 218

1 GlyThr*****ProSer 6
|||||
82 GGCACGAGTGCTCCTCT 65

seq_name: gb_sts:KLAJ9915

seq_documentation_block:
LOCUS KLAJ9915 222 bp DNA linear STS 20-NOV-1998
DEFINITION Kluyveromyces lactis DNA fragment for sequence tagged site, clone
          Okam5f02d.
ACCESSION AJ229915
VERSION AJ229915.1 GI:3820365
KEYWORDS STS.
SOURCE Kluyveromyces lactis.
ORGANISM Kluyveromyces lactis.
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
TITLE Ozier-Kalogeropoulos,O.
JOURNAL Direct Submission
          Submitted (28-MAY-1998) Ozier-Kalogeropoulos O., Institut Pasteur,
          Genetique Moleculaire des Levures25 rue de Dr Roux, 75724 Paris
          cedex 15 FRANCE
REFERENCE 2 (bases 1 to 222)
AUTHORS Ozier-Kalogeropoulos,O., Malpertuy,A., Boyer,J., Tekala,F. and
          Dujon,B.
TITLE Random exploration of the Kluyveromyces lactis genome and
          comparison with that of Saccharomyces cerevisiae
JOURNAL Nucleic Acids Res. 26 (23), 5511-5524 (1998)
MEDLINE 99045620
FEATURES
  source
    1..222
    /organism="Kluyveromyces lactis"
    /strain="CBS2539"
    /db_xref="taxon:28985"
    /clone="Okam5f02d"
    <1..>222
  STS
    69 a 42 c 44 g 67 t
  BASE COUNT
  ORIGIN

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x KLAJ9915

  Align seg 1/1 to: KLAJ9915 from: 1 to: 222

1 GlyThr*****ProSer 6
|||||
157 GGCACCACTCACCCTCT 174

seq_name: gb_pl:AY022702

seq_documentation_block:
LOCUS AY022702 224 bp DNA linear PLN 07-FEB-2001
DEFINITION Oryza sativa microsatellite MRG5027 containing (CCG)X8, genomic
          sequence.
ACCESSION AY022702
VERSION AY022702.1 GI:12705993
KEYWORDS STS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
TITLE Ehrhartoideae; Oryzeae; Oryza.
JOURNAL 1 (bases 1 to 224)
          Simple sequence repeats from Monsanto rice genomic sequences
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
JOURNAL Direct Submission
          Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh
          Blvd., Creve Coeur, MO 63167, USA
          Derived from rice genomic sequences generated from the Monsanto
          Rice Genome Sequencing project. Please see
          http://www.rice-research.org for more information. The sequence

```

```

VERSION AY022702.1 GI:12705918
KEYWORDS STS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
TITLE Ehrhartoideae; Oryzeae; Oryza.
JOURNAL 1 (bases 1 to 224)
          Simple sequence repeats from Monsanto rice genomic sequences
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
JOURNAL Direct Submission
          Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh
          Blvd., Creve Coeur, MO 63167, USA
          Derived from rice genomic sequences generated from the Monsanto
          Rice Genome Sequencing project. Please see
          http://www.rice-research.org for more information. The sequence
          data were produced primarily in the laboratories of Dr. Leroy Hood
          at the University of Washington in Seattle.
FEATURES
  source
    1..224
    /organism="Oryza sativa"
    /db_xref="taxon:4530"
  repeat_region
    1..224
    /note="microsatellite MRG5027"
    /rpt_type=tandem
    /rpt_unit=ccg
  BASE COUNT
  ORIGIN
    25 a 86 c 76 g 37 t

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AY022702

  Align seg 1/1 to: AY022702 from: 1 to: 224

1 GlyThr*****ProSer 6
|||||
197 GGCACCTCGTCGCGCTCG 214

seq_name: gb_pl:AY022777

seq_documentation_block:
LOCUS AY022777 224 bp DNA linear PLN 07-FEB-2001
DEFINITION Oryza sativa microsatellite MRG5102 containing (CGC)X8, closest to
          marker pTA248, genomic sequence.
ACCESSION AY022777
VERSION AY022777.1 GI:12705993
KEYWORDS STS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
TITLE Ehrhartoideae; Oryzeae; Oryza.
JOURNAL 1 (bases 1 to 224)
          Simple sequence repeats from Monsanto rice genomic sequences
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
JOURNAL Direct Submission
          Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh
          Blvd., Creve Coeur, MO 63167, USA
          Derived from rice genomic sequences generated from the Monsanto
          Rice Genome Sequencing project. Please see
          http://www.rice-research.org for more information. The sequence

```

data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.

```

FEATURES
  source
    1. .224
      /organism="Oryza sativa"
      /db_xref="taxon:4530"
  repeat_region
    1. .224
      /note="microsatellite MRG5102"
      /rpt_type=tandem
      /rpt_unit=cgc
  BASE COUNT      32 a   79 c   85 g   28 t
  ORIGIN

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AY022777 ..
  Align seg 1/1 to: AY022777 from: 1 to: 224

    1 GlyThr*****ProSer 6
      |||||||
    58 GGAATTGGCGCCGCTCA 75

seq_name: gb_pl:AY023473

seq_documentation_block:
  LOCUS      AY023473      224 bp      DNA      linear      PLN 07-FEB-2001
  DEFINITION Oryza sativa microsatellite MRG5798 containing (TCT)X8, genomic
  sequence.
  VERSION     AY023473
  KEYWORDS    Oryza sativa.
  SOURCE      Oryza sativa.
  ORGANISM    Oryza sativa.
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
  REFERENCE   1 (bases 1 to 224)
  AUTHORS     Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
  TITLE       Simple sequence repeats from Monsanto rice genomic sequences
  JOURNAL     Unpublished
  REFERENCE   2 (bases 1 to 224)
  AUTHORS     Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
  TITLE       Direct Submission
  JOURNAL     Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh
  COMMENT     Blvd., Creve Coeur, MO 63167, USA
              Derived from rice genomic sequences generated from the Monsanto
              Rice Genome Sequencing project. Please see
              http://www.rice-research.org for more information. The sequence
              data were produced primarily in the laboratories of Dr. Leroy Hood
              at the University of Washington in Seattle.
FEATURES
  source
    1. .224
      /organism="Oryza sativa"
      /db_xref="taxon:4530"
  repeat_region
    1. .224
      /note="microsatellite MRG5798"
      /rpt_type=tandem
      /rpt_unit=tct
  BASE COUNT      32 a   82 c   67 g   43 t
  ORIGIN

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AY023473 rev ..
  Align seg 1/1 to reverse of: AY023473 from: 1 to: 224

    1 GlyThr*****ProSer 6
      |||||||
    146 GGTACGGCTGGCGCTCG 129

seq_name: gb_ov:AF005431

seq_documentation_block:
  LOCUS      AF005431      226 bp      DNA      linear      VFT 04-AUG-1998
  DEFINITION Varanus komodoensis microsatellite sequence.
  ACCESSION   AF005431
  VERSION     AF005431.1 GI:3378037
  KEYWORDS    Varanus komodoensis.
  SOURCE      Varanus komodoensis.
  ORGANISM    Varanus komodoensis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae;
  Varanus.
  REFERENCE   1 (bases 1 to 226)
  AUTHORS     Ciofi,C. and Bruford,M.W.
  TITLE       Isolation and characterization of microsatellite loci in the Komodo
  JOURNAL     dragon Varanus komodoensis
  REFERENCE   2 (bases 1 to 226)
  AUTHORS     Ciofi,C. and Bruford,M.W.
  TITLE       Direct Submission
  JOURNAL     Submitted (28-MAY-1997) Conservation Genetics Group, Institute of
  FEATURES     Zoology, Regent's Park, London NW1 4RY, UK
  SOURCE      Location/Qualifiers
    1. .226
      /organism="Varanus komodoensis"
      /db_xref="taxon:61221"
      /note="microsatellite"
      /rpt_type=tandem
      /rpt_unit=ac
  BASE COUNT      54 a   70 c   54 g   48 t
  ORIGIN

alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AF005431 ..
  Align seg 1/1 to: AF005431 from: 1 to: 226

    1 GlyThr*****ProSer 6
      |||||||
    17 GGTACCGCAGCTCTTCT 34

seq_name: gb_pl:PHVCHS

seq_documentation_block:
  LOCUS      PHVCHS      226 bp      mRNA      linear      PLN 27-APR-1993
  DEFINITION Phaseolus vulgaris chalcone synthase mRNA.
  ACCESSION   K02953
  VERSION     K02953.1 GI:169334
  KEYWORDS    chalcone synthetase; synthetase.
  SOURCE      Phaseolus vulgaris L. (common bean), cDNA to mRNA, clone pCHS1.
  ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Phaseolus.
  REFERENCE   1 (bases 1 to 226)

```

AUTHORS Ryder,T.B., Cramer,C.L., Bell,J.N., Robbins,M.P., Dixon,R.A. and Lamb,C.J.
TITLE Elicitor rapidly induces chalcone synthase mRNA in Phaseolus vulgaris cells at the onset of the phytoalexin defense response
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81, 5724-5728 (1984)
COMMENT Draft entry and sequence for [1] kindly provided by T.Ryder in computer-readable form, 12-JAN-1985.
FEATURES
 source Location/Qualifiers
 1..226
 /organism="Phaseolus vulgaris"
 /db_xref="taxon:3885"
 <1..>226
 /notes="chalcone synthase"
 /codon_start=1
 /protein_id="AAA33758.1"
 /db_xref="GI:553091"
 /translation="EIROQRAEGPATILAIQTATPNCVDOSTYDPYVFRITNSEHM
 TDLKEKFORMKDKMKIKRYMHLDEILKEI"
BASE COUNT 78 a 55 c 52 g 41 t
ORIGIN 110 bp upstream of HinfI site.

alignment_scores:
 Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x PHVCHS ..

Align seg 1/1 to: PHVCHS from: 1 to: 226

1 GlyThr*****ProSer 6
|||||
52 GGACTGCAACCCCATCA 69

seq_name: gb_pr:HS49E3F

seq_documentation_block:
LOCUS HS49E3F 230 bp DNA linear PRI 23-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 49e3, forward
 read cpq49e3.ft1a.
ACCESSION 265624
VERSION 265624.1 GI:1038446
KEYWORDS CpG island; genomic MseI fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 230)
AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
REFERENCE 2 (bases 1 to 230)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
 source Location/Qualifiers
 1..230
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /sex="male"
 /dev_stage="adult"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /clone="49e3"

BASE COUNT 70 a 50 c 53 g 55 t 2 others
ORIGIN

alignment_scores:
 Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x HS49E3F/rev ..

Align seg 1/1 to reverse of: HS49E3F from: 1 to: 230

1 GlyThr*****ProSer 6
|||||
121 GGACATCAAGTCCCTCA 104

alignment_scores: Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAF96301 ..

Align seg 1/1 to: AAF96301 from: 1 to: 21

1 GlyThr*****ProSer 6
||||| 6
4 GGCACCTACGTCGCCCTCC 21

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAC83198

seq_documentation_block:
ID AAC83198 standard; DNA; 21 BP.
XX AC AAC83198;
XX DT 02-MAR-2001 (first entry)
XX DE PCR primer R1 used in growth hormone gene sequence identification.
XX KW Ostrich; growth hormone; transgenic bird; growth promotion;
XX KW PCR primer; ss.
XX OS Synthetic.
XX PN JP2000270870-A.
XX PD 03-OCT-2000.
XX PF 24-MAR-1999; 99JP-0079699.
XX PR 24-MAR-1999; 99JP-0079699.
XX PA (UYKI-) UNIV KINKI.
XX PA (NIGA-) NIPPON GAKUJUTSU SHINKOKAI.
XX DR WPI; 2001-027419/04.
XX PT An ostrich growth hormone polypeptide, a gene and a transformed bird
XX PS Example; Page 4; 8pp; Japanese.
XX CC This invention relates to an ostrich derived growth hormone gene and
CC protein sequences AAC83185 and AAB36000. The invention includes fragments
CC of the growth hormone protein which retain their growth promoting
CC activity. Included in the invention is a recombinant plasmid containing
CC the gene, a method for introducing the gene in to a bird, and a
CC transformed bird produced by the method. The method is used for promoting
CC the growth of an animal. This sequence represents a primer used in the
CC isolation and characterisation of the ostrich growth hormone gene of the
CC invention.
SQ Sequence 21 BP; 4 A; 9 C; 3 G; 5 T; 0 other;

alignment_scores: Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAC83198 ..

Align seg 1/1 to: AAC83198 from: 1 to: 21

1 GlyThr*****ProSer 6
||||| 6
2 GGCACCTCATCACCTTCC 19

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAS96743

seq_documentation_block:
ID AAS96743 standard; DNA; 21 BP.
XX AC AAS96743;
XX DT 26-FEB-2002 (first entry)
XX DE Demeter gene PCR primer SKES-4.
XX DE Deneter; DMT; Atropos; ATR; 5-methylcytosine glycosylase; ss;
XX KW DNA demethylation; transgenic plant; transcription modulation;
XX KW flowering time; endosperm development; MEDEA; PCR primer.
XX OS Planta.
XX PN WO200180626-A1.
XX PD 01-NOV-2001.
XX PF 23-APR-2001; 2001WO-US13059.
XX PR 21-APR-2000; 2000US-0553690.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Fischer RL, Choi Y, Hannon M, Okumuro JK, Tatarinova TV;
XX WPI; 2002-055307/07.
XX PT New polynucleotide that control plant development comprising a sequence
XX having a specific homology to DEMETER domains A,B or C
XX Disclosure; Page 24; 109pp; English.
XX CC The invention relates to an isolated polynucleotide sequence or their
CC complement encoding a polypeptide having a sequence at least 40%
CC identical to DMT (DEMETER, previously known as ATROPOS (ATR)) Domain A, B
CC or C or their combinations. Also included are an expression
CC cassette comprising the polynucleotide or comprising a heterologous
CC polynucleotide under the control of a promoter at least 70% identical to
CC DMT 5' flanking sequence, DMT 3' flanking sequence or an 5' untranslated
CC region of DMT, a host cell comprising an exogenous polynucleotide
CC encoding a DMT-like protein and a transgenic plant comprising a
CC polynucleotide encoding a DMT-like protein. The expression cassette is
CC useful for modulating transcription. The method comprises introducing
CC the cassette into a host cell preferably Agrobacterium by sexual
CC cross, and selecting a host cell with modulated transcription, where
CC the protein is capable of exhibiting at least one of the following
CC biological activities, which include enhanced expression of the
CC protein in a plant results in a delay in flowering time, introduction of
CC the protein into a cell results in modulation of methylation of
CC chromosomal DNA in the cell, reduction of expression of the protein in a
CC plant results in enhanced endosperm development and expression of the
CC protein in an Arabidopsis leaf results in expression of the MEDEA gene.
CC The polynucleotide is useful for detecting a nucleic acid in a sample.
CC DEMETER is related to 5-methylcytosine glycosylases and regulates
CC transcription of target genes by demethylation. The present
CC sequence represents a PCR primer used to isolate the nucleic acids
CC encoding the DMT-like proteins of the invention.
XX
SQ Sequence 21 BP; 6 A; 7 C; 5 G; 3 T; 0 other;

alignment_scores: Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAS96743 ..

Align seg 1/1 to: AAS96743 from: 1 to: 21

1 GlyThr*****ProSer 6
|||||
2 GGAACAAGTCACCATCT 19

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:AAS96765

seq_documentation_block:

ID AAS96765 standard; DNA; 21 BP.

XX AC AAS96765;

DT 26-FEB-2002 (first entry)

DE Demeter gene PCR primer SKES-4.

XX Demeter; DMT; Atropis; ATR; 5-methylcytosine glycosylase; ss;
KW DNA demethylation; transgenic plant; transcription modulation;
KW flowering time; endosperm development; MEDEA; PCR primer.

XX Planta.

XX WO200180626-A1.

XX 01-NOV-2001.

XX 23-APR-2001; 2001WO-US13059.

XX 21-APR-2000; 2000US-0553690.

XX (REGC) UNIV CALIFORNIA.

XX Fischer RL, Choi Y, Hannon M, Okamuro JK, Tatarinova TV;

XX WPI; 2002-055307/07.

PT New polynucleotide that control plant development comprising a sequence
PT having a specific homology to DEMETER domains A,B or C -

PS Disclosure; Page 40; 109pp; English.

XX The invention relates to an isolated polynucleotide sequence or their
CC complement encoding a polypeptide having a sequence at least 40%
CC identical to DMT (DEMETER, previously known as APROPOS (ATR)) Domain A, B
CC or C or their combinations. Also included are an expression
CC cassette comprising the polynucleotide or comprising a heterologous
CC polynucleotide under the control of a promoter at least 70% identical to
CC DMT 5' flanking sequence, DMT 3' flanking sequence or an 5' untranslated
CC region of DMT, a host cell comprising an exogenous polynucleotide
CC encoding a DMT-like protein and a transgenic plant comprising a
CC polynucleotide encoding a DMT-like protein. The expression cassette is
CC useful for modulating transcription. The method comprises introducing
CC the cassette into a host cell preferably Agrobacterium by sexual
CC cross, and selecting a host cell with modulated transcription, where
CC the protein is capable of exhibiting at least one of the following
CC biological activities, which include enhanced expression of the
CC protein in a plant results in a delay in flowering time, introduction of
CC the protein into a cell results in modulation of methylation of
CC chromosomal DNA in the cell, reduction of expression of the protein in a
CC plant results in enhanced endosperm development and expressing of the
CC protein in an Arabidopsis leaf results in expression of the MEDEA gene.
CC The polynucleotide is useful for detecting a nucleic acid in a sample.
CC DEMETER is related to 5-methylcytosine glycosylases and regulates
CC transcription of target genes by demethylation. The present
CC sequence represents a PCR primer used to isolate the nucleic acids
CC encoding the DMT-like proteins of the invention.

XX SQ Sequence 21 BP; 6 A; 7 C; 5 G; 3 T; 0 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAS96765 ..

Align seg 1/1 to: AAS96765 from: 1 to: 21

1 GlyThr*****ProSer 6
|||||
2 GGAACAAGTCACCATCT 19

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAS36773

seq_documentation_block:

ID AAZ36773 standard; DNA; 23 BP.

XX AC AAZ36773;

DT 13-MAR-2000 (first entry)

DE Oligonucleotide probe/primer AHCPE6 derived from the AHCP gene.

XX Human; AHCP gene; autosomal highly conserved protein; schizophrenia;
KW neurological disease; genetic predisposition; chromosome 6p23; D6S274;
KW D6S285; psychological disease; gene therapy; probe; primer; ss.

XX Synthetic.

XX Homo sapiens.

XX WO9957316-A1.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-IB00846.

XX 30-APR-1998; 98US-0083625.

XX 31-DEC-1998; 98US-0114592.

XX (INSP) INST PASTEUR.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Leroy P, Bourgeron T, McElreavey K, Fellous M, Jamain S;

XX WPI; 2000-086415/07.

XX New gene encoding autosomal high conserved protein used to diagnose a
XX genetic predisposition to schizophrenia -

XX Claim 7; Page 11; 76pp; English.

XX Oligonucleotides AAZ36758-76 are derived from the human AHCP (autosomal
CC highly conserved protein) gene. The oligonucleotides are useful as
CC probes and primers. The AHCP gene is linked to a genetic
CC predisposition to schizophrenia. The gene is located on chromosome
CC 6p23, between markers D6S274 and D6S285. Several polymorphisms
CC are found in the AHCP gene. Oligonucleotide probes derived from the
CC AHCP sequences can be used to screen for patients having a genetic
CC predisposition for a neurological or psychological disease, especially
CC schizophrenia. The invention is used to diagnose a genetic
CC predisposition to schizophrenia, and to treat the disorder by gene
CC therapy. The invention provides a treatment that is specific to
CC schizophrenic disorders, without the risk of significant side effects.

XX SQ Sequence 23 BP; 3 A; 6 C; 11 G; 3 T; 0 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:
BASK-853-CLAIM5 x AAZ36773/rev ..
Align seg 1/1 to reverse of: AAZ36773 from: 1 to: 23
1 GlyThr*****ProSer 6
|||||
23 GGGACCACGAGCGCGTCC 6

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAZ36744

seq_documentation_block:
ID AAS96744 standard; DNA; 29 BP.
XX
AC AAS96744;
XX
DT 26-FEB-2002 (first entry)
XX
DE Demeter gene PCR primer SKEN-6.
XX
KW Demeter; DMT; Atropos; ATR; 5-methylcytosine glycosylase; ss;
KW DNA demethylation; transgenic plant; transcription modulation;
KW flowering time; endosperm development; MEDEA; PCR primer.
XX
OS Planta.
XX
PN WO200180626-A1.
XX
PD 01-NOV-2001.
XX
PF 23-APR-2001; 2001WO-US13059.
XX
PR 21-APR-2000; 2000US-0553690.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Fischer RL, Choi Y, Hannon M, Okamuro JK, Tatarinova TV;
XX WPI; 2002-055307/07.
XX
New polynucleotide that control plant development comprising a sequence
having a specific homology to DEMETER domains A,B or C
Disclosure; Page 24; 109pp; English.
XX
The invention relates to an isolated polynucleotide sequence or their
complement encoding a polypeptide having a sequence at least 40%
identical to DMT (DEMETER, previously known as ATROPOS (ATR)) Domain A, B
or C or their combinations. Also included are an expression
cassette comprising the polynucleotide or comprising a heterologous
polynucleotide under the control of a promoter at least 70% identical to
DMT 5' flanking sequence, DMT 3' flanking sequence or an 5' untranslated
region of DMT, a host cell comprising an exogenous polynucleotide
encoding a DMT-like protein and a transgenic plant comprising a
polynucleotide encoding a DMT-like protein. The expression cassette is
useful for modulating transcription. The method comprises introducing
the cassette into a host cell preferably Agrobacterium by sexual
cross, and selecting a host cell with modulated transcription, where
the protein is capable of exhibiting at least one of the following
biological activities, which include enhanced expression of the
protein in a plant results in a delay in flowering time, introduction of
the protein into a cell results in modulation of methylation of
chromosomal DNA in the cell, reduction of expression of the protein in a
plant results in enhanced endosperm development and expressing of the
protein in an Arabidopsis leaf results in expression of the MEDEA gene.
The polynucleotide is useful for detecting a nucleic acid in a sample.
DEMETER is related to 5-methylcytosine glycosylases and regulates
transcription of target genes by demethylation. The present
sequence represents a PCR primer used to isolate the nucleic acids
encoding the DMT-like proteins of the invention.
XX
Sequence 29 BP; 8 A; 9 C; 6 G; 6 T; 0 other;

```

```

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAS96744 ..
Align seg 1/1 to: AAS96744 from: 1 to: 29
1 GlyThr*****ProSer 6
|||||
11 GGACCAAGTGCACCATCT 28

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAZ38959

seq_documentation_block:
ID AAT38959 standard; DNA; 34 BP.
XX
AC AAT38959;
XX
DT 20-MAY-1997 (first entry)
XX
DE Mouse alpha-1,3-galactosyltransferase gene fragment B PCR primer.
XX
KW Primer; polymerase chain reaction; PCR; amplification; fragment B;
KW mouse; murine; alpha-1,3-galactosyltransferase; construction;
KW plasmid; pGT/FT-2; human; alpha-1,2-fucosyltransferase; H-type;
KW antigen; non-primate; mammal; pig; tissue; transgenic; sugar;
KW transferase; transplant; hyperacute rejection; ss.
XX
OS Synthetic.
XX
PN WO9628967-A1.
XX
PD 26-SEP-1996.
XX
PF 18-MAR-1996; 96WO-JP00703.
XX
PR 17-MAR-1995; 95WO-JP00488.
XX
PA (KOIK/) KOIKE C.
XX
PI Koike C;
XX
WPI; 1996-442870/44.
XX
Transgenic non-primate mammal contg. alpha-1,2-fucosyltransferase
of higher primate - expresses H-type antigen and is source of
compatible tissue for transplantation into higher primate
Example 9; Page 42; 65pp; Japanese.
XX
The present sequence is a primer for the PCR amplification of
fragment B of the mouse alpha-1,3-galactosyltransferase (GT) gene,
which was used in the construction of the plasmid pGT/FT-2.
CC pGT/FT-2, which contains the murine GT and human
CC alpha-1,2-fucosyltransferase genes, was injected into the
CC fertilised ova of a non-primate mammal, e.g. a pig. Tissues from
CC the resultant transgenic mammals, which contain the human sugar
CC transferase gene and express a H-type antigen, can be
CC transplanted into a human without inducing hyperacute rejection.
XX
Sequence 34 BP; 6 A; 13 C; 7 G; 8 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

```



```
BASK-853-CLAIM5 x AAT38959
Align seg 1/1 to: AAT38959 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACACCAGTCCTTCT 26

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV35870
seq_documentation_block:
ID AAV35870 standard; DNA; 34 BP.
XX
AC AAV35870;
XX
XX
XX 26-AUG-1998 (first entry)
XX
XX PCR primer GT4 of the invention.
XX
XX DNA construct; homologous recombination; cell surface substance;
XX PCR primer; ss.
XX
XX Synthetic.
XX
XX JF10146193-A.
XX
XX 02-JUN-1998.
XX
XX 19-APR-1997; 97JP-0116372.
XX
XX 23-SEP-1996; 96JP-0287242.
XX
XX (KOIK/) KOIKE C.
XX
XX WPI; 1998-369867/32.
XX
XX New DNA construct for homologous recombination - and selection of
XX transformant
XX
XX Example 1; Page 11; 17pp; Japanese.
XX
XX PCR primers AAV35869-70 represent primers used to exemplify the
XX invention. The specification describes DNA constructs for homologous
XX recombination. The construct contains a sequence homologous to the DNA
XX sequence of a host cell synthesising a cell surface substance and a DNA
XX sequence encoding the protein participating in the synthesis of the cell
XX surface substance. The protein is constituted so as to be expressed and
XX controlled by the expression control region of the protein
XX participating in the synthesis of the cell surface substance of the
XX reception cell. The DNA construct can be used to solve the problems in
XX conventional homologous recombination.
XX
XX Sequence 34 BP; 7 A; 12 C; 7 G; 8 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAV35870
Align seg 1/1 to: AAV35870 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACACCAGTCCTTCT 26

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV27023
seq_documentation_block:
ID AAV27023 standard; DNA; 34 BP.
XX
AC AAV27023;
XX
XX 01-SEP-1998 (first entry)
XX
XX Primer GT4.
XX
XX ss; transgenic animal; embryonic stem cell; PCR; primer; amplification.
XX
XX Synthetic.
XX
XX Mus sp.
XX
XX JF10150990-A.
XX
XX 09-JUN-1998.
XX
XX 18-NOV-1996; 96JP-0322352.
XX
XX 18-NOV-1996; 96JP-0322352.
XX
XX (KOIK/) KOIKE C.
XX
XX WPI; 1998-379993/33.
XX
XX Introducing heterologous gene - and preparation of non-human
XX transgenic animal
XX
XX Example -; Page 5; 6pp; Japanese.
XX
XX The primers AAV27020-V27025 were used in the method of the invention to
XX introduce a heterologous gene to a cell creating a non-human
XX transgenic animal. The gene was introduced by transplanting an
XX embryonic stem cell containing the desired DNA sequence into a
XX fertilised egg. This egg was then transplanted into the uterus of a
XX pseudo-pregnant non-human animal.
XX
XX Sequence 34 BP; 7 A; 12 C; 7 G; 8 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAV27023
Align seg 1/1 to: AAV27023 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACACCAGTCCTTCT 26

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV99851
seq_documentation_block:
ID AAV99851 standard; DNA; 34 BP.
XX
AC AAV99851;
XX
XX 30-SEP-1999 (first entry)
XX
XX PCR primer GT4 for mouse PGK gene.
XX
XX PCR primer; transgenic animal; mouse; phosphoglucokinase; PGK; ss.
XX
XX Synthetic.
XX
XX Mus sp.
XX
XX JF11192036-A.
XX
XX 21-JUL-1999.
XX
```

```
PF 04-JAN-1998; 98JP-0030322.
PR 04-JAN-1998; 98JP-0030322.
XX (KOIK/) KOIKE C.
XX WPI; 1999-461420/39.
XX New process - useful for preparation of transgenic animals
XX
XX Example 2; Page 5; 6pp; Japanese.
XX
XX This sequence represents a PCR primer for the mouse phosphoglucokinase
XX (P6K) gene. The invention relates to a process for the preparation of
XX transgenic animals, that comprises: (1) introduction of a foreign gene
XX into foetal cells; (2) screening of the recombinant cells from the foetal
XX cells; (3) injection of the screened recombinant cells into fertilised
XX eggs; and (4) transplantation of the injected fertilised eggs into the
XX uterus of pseudopregnant female animals. The method is used to prepare
XX homologous recombinant transgenic animals.
XX
XX Sequence 34 BP; 7 A; 12 C; 7 G; 8 T; 0 other;
SQ
alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AAX99851 ..
Align seg 1/1 to: AAX99851 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACCAACCACTCTCTCT 26
seq_name: /SDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAS05363
seq_documentation_block:
ID AAS05363 standard; DNA; 34 BP.
XX
AC AAS05363;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mouse alpha-1,3 galactosyltransferase exon 5 and exon 6 primer #2.
XX
DE Mouse; murine; alpha-1,3 galactosyltransferase promoter; primer;
KW recombinant expression cassette; gene expression; transgenic animal;
KW livestock; growth hormone; xenotransplantation; transgenic cell; ss.
XX
OS Mus musculus.
XX
PN WO200130992-A2.
XX
PD 03-MAY-2001.
XX
PF 20-OCT-2000; 2000WO-US29139.
XX
PR 22-OCT-1999; 99US-0161092.
PR 25-AUG-2000; 2000US-0227951.
XX
XX (UYPI-) UNIV PITTSBURGH.
XX
XX Koike C;
XX
XX WPI; 2001-300505/31.
XX
XX New recombinant expression cassette useful in xenotransplantation -
XX
XX Disclosure; Page 129; 138pp; English.
```

```
XX
XX The present sequence for mouse alpha-1,3 galactosyltransferase (GT)
XX exon 5 and exon 6 primer #2 is used for identifying this region
XX of the alpha-1,3 GT nucleotide sequence. Various alpha-1,3 GT gene
XX sequences from different animals (AAS05290-AAS05333) and the primers
XX (AAS05334-AAS05385) used to isolate these sequences are given. The
XX invention provides a novel recombinant expression cassette comprising
XX an alpha-1,3 GT promoter operably linked to a polynucleotide for
XX expression. The cassettes of the invention can be used to express
XX foreign genes or to disrupt the native alpha-1,3 GT genomic sequence.
XX The invention provides transgenic animals and methods for their
XX production. Transgenic mice in which a reporter gene is operably linked
XX to the alpha-1,3 GT promoter can be used for assessing promoter activity
XX and specificity. Transgenic livestock having the novel expression
XX cassette in which a growth hormone is expressed under the control of the
XX alpha-1,3 GT promoter can be matured or grown better than commonly
XX employed strains. Tissue obtained from transgenic animals can be
XX implanted into a host providing a method of xenotransplantation from a
XX transgenic animal. The invention also provides a transgenic organ
XX consisting essentially of transgenic cells engineered, for use in
XX transplantation. The methods of the invention facilitate
XX xenotransplantation between species, particularly between species
XX exhibiting differential expression of the gal-alpha-gal epitope.
XX
XX Sequence 34 BP; 6 A; 13 C; 7 G; 8 T; 0 other;
SQ
alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AAS05363 ..
Align seg 1/1 to: AAS05363 from: 1 to: 34
1 GlyThr*****ProSer 6
|||||
9 GGAACCAACCACTCTCTCT 26
seq_name: /SDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI66111
seq_documentation_block:
ID AAI66111 standard; DNA; 41 BP.
XX
AC AAI66111;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human NAD-dependent 2-hydroxyl acid dehydrogenase 16 probe 1.
XX
DE Human; NAD-dependent 2-hydroxyl acid dehydrogenase family protein 16;
KW cytostatic; virucidal; immunomodulatory; antiinflammatory; haemostatic;
KW malignant tumour; human immunodeficiency virus; HIV; infection;
KW immunological disease; gene therapy; probe; ss.
XX
OS Homo sapiens.
XX
PN WO200185922-A1.
XX
XX 15-NOV-2001.
PD
XX 26-FEB-2001; 2001WO-CN00178.
PF
XX 15-MAR-2000; 2000CN-0114919.
PR
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
PA
XX Mao Y, Xie Y;
PI
XX WPI; 2001-648987/74.
DR
XX
```

PT New human nicotinamide adenine dinucleotide (NAD)-dependent 2-hydroxyl
PT acid dehydrogenase family protein 16 for diagnosing and treating e.g.
PT malignant tumours, human immunodeficiency virus infection, and
XX inflammation -
XX
PS Example 6; Page 21; 40pp; Chinese.
XX
CC The invention relates to human nicotinamide adenine dinucleotide (NAD)
CC dependent 2-hydroxyl acid dehydrogenase family protein 16 with
CC cytosolic, virucidal, immunomodulatory, antiinflammatory and
CC haemostatic activity. The protein and encoding polynucleotide are used
CC in diagnosis and treatment of malignant tumour, haemopathy, human
CC immunodeficiency virus (HIV) infection, immunological diseases and
CC various inflammations. The polynucleotide is useful in gene therapy.
CC The present sequence is that of a probe, useful to the invention.
XX
SQ Sequence 41 BP; 5 A; 12 C; 17 G; 7 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAI66111 ..
Align seg 1/1 to: AAI66111 from: 1 to: 41
1 GlyThr*****ProSer 6
|||||
22 GGCACAGCCAGCCCTCT 39

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI66112

seq_documentation_block:
ID AAI66112 standard; DNA; 41 BP.
XX
AC AAI66112;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human NAD-dependent 2-hydroxyl acid dehydrogenase 16 probe 2.
XX
KW Human; NAD-dependent 2-hydroxyl acid dehydrogenase family protein 16;
KW cytosolic; virucidal; immunomodulatory; antiinflammatory; haemostatic;
KW malignant tumour; human immunodeficiency virus; HIV; infection;
KW immunological disease; gene therapy; probe; ss.
XX
OS Homo sapiens.
XX
FN WO200185922-A1.
XX
PD 15-NOV-2001.
XX
PF 26-FEB-2001; 2001WO-CN00178.
XX
PR 15-MAR-2000; 2000CN-0114919.
XX
PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-648987/74.
XX
XX New human nicotinamide adenine dinucleotide (NAD)-dependent 2-hydroxyl
PT acid dehydrogenase family protein 16 for diagnosing and treating e.g.
PT malignant tumours, human immunodeficiency virus infection, and
PT inflammation -
XX
XX Example 6; Page 22; 40pp; Chinese.
XX
CC The invention relates to human nicotinamide adenine dinucleotide (NAD)

CC dependent 2-hydroxyl acid dehydrogenase family protein 16 with
CC cytosolic, virucidal, immunomodulatory, antiinflammatory and
CC haemostatic activity. The protein and encoding polynucleotide are used
CC in diagnosis and treatment of malignant tumour, haemopathy, human
CC immunodeficiency virus (HIV) infection, immunological diseases and
CC various inflammations. The polynucleotide is useful in gene therapy.
CC The present sequence is that of a probe, useful to the invention.
XX
SQ Sequence 41 BP; 5 A; 12 C; 17 G; 7 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAI66112 ..
Align seg 1/1 to: AAI66112 from: 1 to: 41
1 GlyThr*****ProSer 6
|||||
22 GGCACAGCCAGCCCTCT 39

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAZ66424

seq_documentation_block:
ID AAZ66424 standard; DNA; 47 BP.
XX
AC AAZ66424;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human map-related biallelic marker SEQ ID NO:771.
XX
KW Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation;
KW diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(24,G)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO9954500-A2.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1999; 99WO-IB00822.
XX
PR 21-APR-1998; 98US-0082614.
PR 23-NOV-1998; 98US-0109732.
XX
PA (GEST) GENSET.
XX
PI Cohen D, Blumenfeld M, Chumakov I;
XX
DR WPI; 2000-013267/01.
XX
PT Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome -
XX
PS Claim 1; Page 395; 2745pp; English.
XX
CC AAZ66554 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the
CC invention have a variety of uses: they can be used for high density

CC mapping of the human genome, and in complex association studies and
CC haplotyping studies which are useful in determining the genetic basis
CC for disease states. Compositions and methods of the invention can also
CC be useful for the identification of the targets for the development of
CC pharmaceutical agents and diagnostic methods, as well as the
CC characterisation of the differential efficacious responses to and side
CC effects from pharmaceutical agents acting on a disease as well as other
CC treatment.
CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC and 3367, are not actually given a sequence in the Sequence Listing
CC from the present invention.
XX
SQ Sequence 47 BP; 15 A; 8 C; 18 G; 6 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAZ66424/rev ..
Align seg 1/1 to reverse of: AAZ66424 from: 1 to: 47

1 GlyThr*****ProSer 6
|||||
23 GGCACCTGGCGCCGCCAGC 6

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ69350

seq_documentation_block:

ID AAZ69350 standard; DNA; 47 BP.

XX AC AAZ69350;

XX DT 10-SEP-2001 (first entry)

XX DE Human map-related biallelic marker SEQ ID NO:3706.

XX KW Human genome; biallelic marker; high density disequilibrium map;
XX KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
XX KW haplotyping; hybridisation; identification; characterisation;
XX KW diagnosis; single nucleotide polymorphism; SNP; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT variation replace(24,G)
XX FT /*tag= a

XX FT /standard_name= "single nucleotide polymorphism"

XX PN WO9954500-A2.

XX PD 28-OCT-1999.

XX PF 21-APR-1999; 99WO-IB00822.

XX PR 21-APR-1998; 98US-0082614.

XX PR 23-NOV-1998; 98US-0109732.

XX PA (GEST) GENSET.

XX PI Cohen D, Blumenfeld M, Chumakov I;

XX DR WPI; 2000-013267/01.

XX PT Novel biallelic markers used to construct a high density disequilibrium
XX PT map of the human genome

XX PS Claim 3; Page 1021; 2745pp; English.

XX CC AAZ65654 to AAZ69578 represent human biallelic markers from the present

CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the
CC invention have a variety of uses; they can be used for high density
CC mapping of the human genome, and in complex association studies and
CC haplotyping studies which are useful in determining the genetic basis
CC for disease states. Compositions and methods of the invention can also
CC be useful for the identification of the targets for the development of
CC pharmaceutical agents and diagnostic methods, as well as the
CC characterisation of the differential efficacious responses to and side
CC effects from pharmaceutical agents acting on a disease as well as other
CC treatment.
CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC and 3367, are not actually given a sequence in the Sequence Listing
CC from the present invention.
XX
SQ Sequence 47 BP; 13 A; 17 C; 8 G; 9 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAZ69350 ..

Align seg 1/1 to: AAZ69350 from: 1 to: 47

1 GlyThr*****ProSer 6
|||||

2 GGAACAGCCAGCCCATCT 19

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAZ696399

seq_documentation_block:

ID AAZ696399 standard; DNA; 49 BP.

XX AC AAZ696399;

XX DT 26-FEB-2002 (first entry)

XX DE Arabidopsis antisense PCR primer for cell cycle protein CCP35.

XX KW Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;
XX KW plant growth regulator; plant development; abiotic stress; biotic stress;
XX KW nutrient deprivation; pathogen attack; crop yield; PCR primer.

XX OS Arabidopsis thaliana.

XX PN WO200185946-A2.

XX PD 15-NOV-2001.

XX PF 14-MAY-2001; 2001WO-IB01307.

XX PR 12-MAY-2000; 2000US-204045P.

XX PA (CROP-) CROPDESIGN NV.

XX PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX DR WPI; 2002-062249/08.

XX PT New cell cycle protein and nucleic acid molecule encoding it useful for
XX PT regulating cell cycle progression in plants and for identifying
XX PT modulators which are useful as herbicides or plant growth regulators -
XX PS Example 2; Page 279; 316pp; English.

XX CC The invention relates to a novel cell cycle protein (CCP) and the
XX CC polynucleotides encoding them. CCP is useful for identifying a compound
XX CC which modulates the activity of the polypeptide and which binds to the

CC polypeptide and an anti-CCP antibody is useful for detecting the presence
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.
 CC CCP nucleic acid and polypeptide molecules are useful as modulating
 CC agents in regulating cell cycle progression in plants. CCP is useful to
 CC treat disorders characterised by insufficient or excessive production of
 CC CCP protein or production of CCP protein forms which have decreased or
 CC aberrant activity. Compounds that bind to or modulate the activity
 CC of CCP polypeptide are useful as herbicides or plant growth regulators.
 CC The polynucleotide is useful for modifying cell fate, plant development,
 CC plant morphology, biochemistry and/or physiology, the length of the G1,
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
 CC stimulation or enhancement of cell division, DNA replication, seed set,
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
 CC initiation and/or development, nodule function, dwarfism in plants,
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
 CC and the anti-CCP antibody are useful in agriculture to modulate the
 CC protein levels or activity of a protein involved in the cell cycle due
 CC to environmental conditions, including abiotic stress such as
 CC cold, nutrient deprivation, heat, drought, salt stress, or biotic
 CC stress such as pathogen attack, to modulate e.g. enhance crop yields,
 CC and attenuate plant architecture, plant quality traits, plant
 CC reproduction and seed development, endoreduplication in storage cells,
 CC storage tissues and/or storage organs of plants or its parts. CCP is
 CC useful as an immunogen to generate antibodies. CCP protein is useful to
 CC screen for naturally occurring CCP substrates. The polynucleotide is
 CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic
 CC lesion in a CCP gene and to modulate CCP activity. The present sequence
 CC is a PCR primer used to isolate a nucleic acid encoding a CCP protein of
 CC the invention.

XX
 SQ Sequence 49 BP; 9 A; 12 C; 14 G; 14 T; 0 other;

alignment_scores:
 Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAS96399/rev ..

Align seg 1/1 to reverse of: AAS96399 from: 1 to: 49

1 GlyThr*****ProSer 6
 |||||
 40 GGACACACGACACCCAGC 23

seq_name: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:AAL30975

seq_documentation_block:

ID AAL30975 standard; DNA; 50 BP.

AC AAL30975;

24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #4183.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.

XX Homo sapiens.

OS

XX WO200147944-A2.

PN

XX

PD 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

PA Shimkets RA, Leach M;

XX WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

PT oncogenes and histones, useful for diagnosing and treating, e.g.

PT cancer, autoimmune diseases and infections -

XX Claim 1; Page 2588; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic

CC variants of proteins related to amylases, amyloid proteins, angiotensin,

CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,

CC histones, kinases, colony stimulating factors, complement related

CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,

CC G-protein coupled receptors and thioesterases. The present sequence is

CC one such oligonucleotide. The oligonucleotides and the peptides encoded

CC by them may be used in the prevention, diagnosis and treatment of

CC diseases associated with inappropriate expression of the proteins listed

CC above. Disorders that may be prevented, diagnosed and/or treated include

CC multifactorial diseases with a genetic component, such as autoimmune

CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,

CC systemic lupus erythematosus and Grave's disease), inflammation, cancer

CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

CC leukaemia), diseases of the nervous system and an infection of pathogenic

CC organisms.

XX

SQ Sequence 50 BP; 9 A; 16 C; 17 G; 8 T; 0 other;

alignment_scores:
 Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAL30975 ..

Align seg 1/1 to: AAL30975 from: 1 to: 50

1 GlyThr*****ProSer 6

|||||

19 GGCACTGCACGCCATCC 36

seq_name: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:AAL30976

seq_documentation_block:

ID AAL30976 standard; DNA; 50 BP.

AC AAL30976;

24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #4184.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.

OS Homo sapiens.
XX WO200147944-A2.
PN 05-JUL-2001.
XX 28-DEC-2000; 2000WO-US35498.
XX 28-DEC-1999; 99US-0173419.
XX 27-DEC-2000; 2000US-0173419.
XX (CURA-) CURAGEN CORP.
XX Shinkets RA, Leach M;
PI WPI; 2001-465210/50.
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX oncogenes and histones, useful for diagnosing and treating, e.g.
XX cancer, autoimmune diseases and infections -
XX Claim 1; Page 2588; 4143pp; English.
XX The present invention relates to oligonucleotides encoding polymorphic
XX variants of proteins related to amylases, amyloid proteins, angiotensin,
XX apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
XX histones, kinases, colony stimulating factors, complement related
XX proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
XX G-protein coupled receptors and thioesterases. The present sequence is
XX one such oligonucleotide. The oligonucleotides and the peptides encoded
XX by them may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate expression of the proteins listed
XX above. Disorders that may be prevented, diagnosed and/or treated include
XX multifactorial diseases with a genetic component, such as autoimmune
XX diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
XX systemic lupus erythematosus and Grave's disease), inflammation, cancer
XX (e.g. cancers of the bladder, brain, breast, colon and kidney,
XX leukemia), diseases of the nervous system and an infection of pathogenic
XX organisms.
XX Sequence 50 BP; 10 A; 16 C; 16 G; 8 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AAL30976 ..
Align seg 1/1 to: AAL30976 from: 1 to: 50
1 GlyThr*****ProSer 6
|||||
18 GGCACCTGCACGCCCATCC 35
seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA64786
seq_documentation_block:
ID AAA64786 standard; DNA; 51 BP.
XX
AC AAA64786;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis pmpG gene PCR primer # 3.
XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection;
KW atherosclerosis; coronary heart disease; antibacterial; ss.
XX

OS Chlamydia trachomatis.
XX WO200034483-A2.
PN 15-JUN-2000.
XX 08-DEC-1999; 99WO-US29012.
XX 08-DEC-1998; 98US-0208277.
XX 08-APR-1999; 99US-0288594.
XX 01-OCT-1999; 99US-0410568.
XX 22-OCT-1999; 99US-0426571.
XX (CORI-) CORIXA CORP.
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises
XX amino acid sequence encoded by polynucleotide sequence -
XX Example 1; Page 225; 256pp; English.
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamidia are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and can
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX and infertility. Trachoma due to ocular infection with C. trachomatis is
XX the leading cause of preventable blindness worldwide. C. pneumonia is a
XX major cause of acute respiratory tract infections in humans and is also
XX thought to play a role in the pathogenesis of atherosclerosis and
XX coronary heart disease. The present sequence is a PCR primer used in the
XX present invention.
XX Sequence 51 BP; 16 A; 16 C; 8 G; 11 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AAA64786 ..
Align seg 1/1 to: AAA64786 from: 1 to: 51
1 GlyThr*****ProSer 6
|||||
5 GGTAACCCGATCACCATCA 22
seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH56289
seq_documentation_block:
ID AAH56289 standard; DNA; 51 BP.
XX
AC AAH56289;
XX
DT 05-SEP-2001 (first entry)
XX
DE Chlamydia trachomatis pmpG gene 5' primer.
XX
KW Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Cap1; CTS29;
KW polymorphic membrane protein; pmp; PCR primer; ss.
XX

OS Chlamydia trachomatis.
PN WO200104074-A2.
XX
XX
XX PD 07-JUN-2001.
XX
XX PF 04-DEC-2000; 2000WO-US32919.
XX
XX PR 03-DEC-1999; 99US-0454584.
XX PR 19-APR-2000; 2000US-0556877.
XX PR 20-JUN-2000; 2000US-0598419.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX
XX DR WPI; 2001-374831/39.
XX
XX PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma; acute respiratory tract infections,
PT atherosclerosis and heart disease -
XX
XX PS Example 1; Page 70; 295pp; English.
XX
XX CC The present sequence is a primer provided in a specification relating to
CC compounds and methods for the treatment and diagnosis of chlamydial
CC infection. The compounds provided include polypeptides and fusion
CC proteins comprising immunogenic portions of Chlamydia antigens
CC and DNA sequences encoding such polypeptides. They are useful for
CC vaccinating against chlamydial infection, which causes pelvic
CC inflammatory disease, trachoma, acute respiratory tract infections,
CC atherosclerosis and heart disease.
XX
XX SQ Sequence 51 BP; 16 A; 16 C; 8 G; 11 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAH56289 ..

Align seg 1/1 to: AAH56289 from: 1 to: 51

1 GlyThr*****ProSer 6
|||||
5 GGTACCGCATCACCACATCA 22

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT:AAZ95714

seq_documentation_block:

ID AAZ95714 standard; DNA; 58 BP.

XX

XX

XX AC AAZ95714;

XX

XX DT 13-JUN-2000 (first entry)

XX

XX DE Clostridium botulinum toxin PCR primer SEQ ID NO:11.

XX

XX KW Clostridium; expression; toxin; vaccine; cosmetic dermatology;

XX KW Clostridium; neurological disease; PCR primer; ss.

XX OS Clostridium botulinum.

XX OS WO200012728-A1.

XX PN

XX PD 09-MAR-2000.

XX

XX PF 23-AUG-1999; 99WO-US19284.

XX

XX PR 28-AUG-1998; 98US-0143634.

XX
XX PA (PROM-) PROMEGA CORP.
XX
XX PI Zdanovsky AG;
XX
XX DR WPI; 2000-256644/22.
XX
XX PT Host cells expressing rare tRNAs and at least one clostridial protein
XX useful for overproduction of clostridial proteins and toxins -
XX
XX PS Example 2; Fig 1; 69pp; English.
XX
XX CC The present invention describes a host cell containing a recombinant
XX expression vector, where the vector encodes transfer RNAs (tRNAs) that
XX recognise rare codons, and where the host cell expresses at least a
XX fragment of at least one clostridial protein. The host cells are useful
XX for the overproduction of Clostridium toxins. These proteins and toxins
XX are useful in medical and veterinary applications, including vaccine
XX production, as well as treatment of neurological and other diseases and
XX conditions. Clostridial genes are extremely A-T rich and as a
XX consequence contain codons specific for amino acids which are rarely
XX used in E. coli. To overcome the expression problems caused by
XX clostridial genes in E. coli, the invention provides host cells with
XX overexpressed cognate tRNAs for rarely used codons. The growth of the
XX host cells is uninhibited by expression of the tRNAs. Production of at
XX least 2 times more clostridial protein in E. coli cells which
XX co-express the tRNA genes as compared to E. coli cells which do not
XX express the tRNA genes is achieved. The methods also avoid the need for
XX an anaerobic environment for growth, permitting large-scale development
XX of antisera, vaccines and toxoids. The present sequence represents a
XX PCR primer for a Clostridium botulinum toxin, which is used in an
XX example from the present invention.
XX
XX SQ Sequence 58 BP; 21 A; 16 C; 11 G; 10 T; 0 other;

alignment_scores:

Quality: 22.00

Ratio: 3.667

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAZ95714 ..

Align seg 1/1 to: AAZ95714 from: 1 to: 58

1 GlyThr*****ProSer 6
|||||
5 GGGACATCATCACCACATCA 22

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1998.DAT:AAV23584

seq_documentation_block:

ID AAV23584 standard; DNA; 62 BP.

XX

XX AC AAV23584;

XX

XX DT 16-JUL-1998 (first entry)

XX

XX DE Primer for luciferase fusion protein gene construction.

XX

XX KW Firefly luciferase; antibody-luciferase fusion protein; PCR primer; ss.

XX OS Synthetic.

XX OS Luciola cruciata.

XX XX JP09187281-A.

XX XX

XX PD 22-JUL-1997.

XX

XX PF 09-JAN-1996; 96JP-0001812.

XX

XX PR 09-JAN-1996; 96JP-0001812.

```
XX PA (KIKK ) KIKKOMAN CORP.
XX DR WPI; 1998-275089/25.
XX PT Antibody-firefly luciferase fused protein - and related products
XX PT i.e. firefly luciferase fused gene, recombinant DNA and its
XX PT preparation
XX PS Example 2; Page 5; 17pp; Japanese.
XX CC This sequence is a primer used in the construction of the gene encoding
XX CC the fusion protein of the invention. The protein is a antibody-firefly
XX CC luciferase fusion protein, in which an antibody part consisting of a
XX CC peptide having antibody activity is combined with an enzyme part
XX CC consisting of firefly luciferase.
XX SQ Sequence 62 BP; 9 A; 21 C; 22 G; 10 T; 0 other;

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAV23584/rev ..
Align seg 1/1 to reverse of: AAV23584 from: 1 to: 62
    1 GlyThr*****ProSer 6
      |||||||
      23 GGGACACGGCACCCTCT 6

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:ABA72555
seq_documentation_block:
ID ABA72555 standard; DNA; 86 BP.
AC ABA72555;
XX DT 01-FEB-2002 (first entry)
XX DE Human foetal liver single exon nucleic acid probe #20860.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00669.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX CC Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human fetal liver.
XX PS Claim 4; SEQ ID NO 20860; 639pp + sequence listing; English.
```

```
XX CC The invention relates to a single exon nucleic acid probe for
XX CC measuring human gene expression in a sample derived from human foetal
XX CC liver. The single exon nucleic acid probes may be used for predicting,
XX CC measuring and displaying gene expression in samples derived from human
XX CC foetal liver. The present sequence is a single exon nucleic acid
XX CC probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 86 BP; 19 A; 25 C; 23 G; 19 T; 0 other;

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x ABA72555/rev ..
Align seg 1/1 to reverse of: ABA72555 from: 1 to: 86
    1 GlyThr*****ProSer 6
      |||||||
      28 GGTACGACGAGTCCTCA 11

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:ABA38293
seq_documentation_block:
ID ABA38293 standard; DNA; 86 BP.
XX AC ABA38293;
XX DT 23-JAN-2002 (first entry)
XX DE Probe #16759 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00666.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX CC Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 4; SEQ ID No 16759; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
```


CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 86 BP; 19 A; 25 C; 23 G; 19 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x ABA38293/rev ..

Align seg 1/1 to reverse of: ABA38293 from: 1 to: 86

1 GlyThr*****ProSer 6
|||||
28 GGTACGACCACTCTTCA 11

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAK20979

seq_documentation_block:

ID AAK20979 standard; DNA; 86 BP.

XX AAK20979;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 20970.

DE Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -

XX Example 4; SEQ ID NO: 20970; 650pp + Sequence Listing; English.

XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

XX Sequence 86 BP; 19 A; 25 C; 23 G; 19 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAK20979/rev ..

Align seg 1/1 to reverse of: AAK20979 from: 1 to: 86

1 GlyThr*****ProSer 6
|||||
28 GGTACGACCACTCTTCA 11

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAK47132

seq_documentation_block:

ID AAK47132 standard; DNA; 86 BP.

XX AAK47132;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 21689.

DE Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 21689; 658pp + Sequence Listing; English.

XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.

XX Sequence 86 BP; 19 A; 25 C; 23 G; 19 T; 0 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAK47132/rev ..

Align seg 1/1 to reverse of: AAK47132 from: 1 to: 86

1 GlyThr*****ProSer 6
|||||
28 GGTACGACGAGTCCTTCA 11

seq_name: /SIDS1/gcgcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI25852

seq_documentation_block:

ID_AAI25852 standard; DNA; 86 BP.

XX AC AAI25852;

XX AC 12-OCT-2001 (first entry)

XX DE Probe #15785 for gene expression analysis in human cervical cell sample.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer; ss.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US006070.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234587.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human cervical epithelial cells -

XX PS Claim 25; SEQ ID No 15785; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes

XX CC (SENP). The present sequence is one such probe. The SENPs are derived

XX CC from human HeLa cells. The SENPs can be used to produce a single exon

XX CC microarray, which can be used for measuring human gene expression in a

XX CC sample derived from human cervical epithelial cells. By measuring gene

XX CC expression, the probes are therefore useful in grading and/or staging

XX CC of diseases of the cervix, notably cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 86 BP; 19 A; 25 C; 23 G; 19 T; 0 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAI25852/rev ..

Align seg 1/1 to reverse of: AAI25852 from: 1 to: 86

1 GlyThr*****ProSer 6
|||||
28 GGTACGACGAGTCCTTCA 11

seq_name: /SIDS1/gcgcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI52968

seq_documentation_block:

ID_AAI52968 standard; DNA; 86 BP.

XX AC AAI52968;

XX AC 17-OCT-2001 (first entry)

XX DE Probe #21654 used to measure gene expression in human placenta sample.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234587.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-48897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human placenta -

XX PS Claim 25; SEQ ID No 21654; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP).

XX CC The present sequence is one such probe. The probes are useful for

XX CC producing a microarray for predicting, measuring and displaying gene

XX CC expression in samples derived from human placenta. The probes are useful

XX CC for antenatal diagnosis of human genetic disorders.

XX SQ Sequence 86 BP; 19 A; 25 C; 23 G; 19 T; 0 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAI52968/rev ..

Align seg 1/1 to reverse of: AAI52968 from: 1 to: 86

1 GlyThr*****ProSer 6
|||||
28 GGTACGACGAGTCCTTCA 11

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT:AAZ90618

seq_documentation_block:

ID AAZ90618 standard; DNA; 102 BP.

XX AAZ90618;

XX 14-JUN-2000 (first entry)

XX Targeting DNA for evolving scFv CDR H3 region to botulinum neurotoxin.

XX Domain specific gene evolution; DSGE; recombinase; transgenic;

XX neurotoxin; ss.

XX Clostridium botulinum.

XX WO200009755-A2.

XX 24-FEB-2000.

XX 12-AUG-1999; 99WO-US18393.

XX 12-AUG-1998; 98US-0096330.

XX 14-AUG-1998; 98US-0133934.

XX (PANG-) PANGENE CORP.

XX Zarling D, Reddy G, Pati S;

XX WPI; 2000-224374/19.

XX Novel methods for evolving specific protein domains using a library of
XX nucleic acid filaments and a recombinase polypeptide used to create
XX animal and plant models of disease -

XX Disclosure; Fig 5; 48pp; English.

XX The invention provides a novel method of domain specific gene evolution
XX (DSGE) of a target nucleic acid encoding an amino acid of interest, that
XX comprises providing pairs of substantially complementary single-stranded
XX targeting polynucleotides (TPs). The pairs comprise a homology clamp
XX substantially corresponding to or substantially complementary to a
XX predetermined sequence of T encoding a domain of A, the pairs comprising
XX a library of mismatches between the TPs and the sequence and a
XX recombinase, to form a library of altered Ts. The methods can be used in
XX the creation of transgenic organisms, and animal and plant models of
XX disease. Domain targeting can also be used in cells or animals that are
XX diseases or altered, e.g. to identify reversion genes. The methods may
XX also be used in the creation of libraries of altered nucleic acids,
XX which are expressed to produce libraries of altered proteins, which can
XX then be screened for useful proteins. The methods are useful to identify
XX new members of gene families which may be useful in functional genomics
XX studies, as well as identification of new drug targets. The invention
XX provides an efficient method of DSGE that generates maximal diversity
XX and increases the probability of identifying a gene of interest. The
XX methods also avoid multiple subcloning steps. Sequences AAZ90616-621
XX represent TPs for evolving CDR region of scFv to botulinum neurotoxin.

XX Sequence 102 BP; 10 A; 17 C; 20 G; 12 T; 43 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAZ90618

Align seg 1/1 to: AAZ90618 from: 1 to: 102

1 GlyThr*****ProSer 6

|||||

83 GGGACCACGGCACCCTCT 100

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT:AAAC13703

seq_documentation_block:

ID AAC13703 standard; cDNA; 124 BP.

XX AAC13703;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 17778.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 17778; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. NO ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.

XX Sequence 124 BP; 14 A; 53 C; 20 G; 37 T; 0 other;

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAC13703

Align seg 1/1 to: AAC13703 from: 1 to: 124

1 GlyThr*****ProSer 6

|||||

34 GGCACTGCTCTCTCTCC 51

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT:AAAC10571

seq_documentation_block:

```

ID  AAC10571 standard; cDNA; 149 BP.
XX  AC
XX  AAC10571;
XX  DT
XX  06-OCT-2000 (first entry)
XX  DE
XX  Human secreted protein 5' EST, SEQ ID NO: 14646.
XX  KW
XX  Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX  KW
XX  gene therapy; chromosome mapping; ss.
XX  OS
XX  Homo sapiens.
XX  PN
XX  EPI033401-A2.
XX  PD
XX  06-SEP-2000.
XX  PF
XX  21-FEB-2000; 2000EP-0200610.
XX  PR
XX  26-FEB-1999; 99US-0122487.
XX  PA
XX  (GEST ) GENSET.
XX  PI
XX  Dumas Milne Edwards J, Duclert A, Giordano J;
XX  DR
XX  WPI; 2000-500381/45.
XX  PS
XX  Claim 1; SEQ ID 14646; 71pp + CD-ROM; English.
XX  PT
XX  The present sequence is one of a large number of 5' ESTs derived from
XX  CC
XX  mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX  CC
XX  identified within the present sequence. The 5' ESTs were prepared from
XX  CC
XX  total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX  CC
XX  sequences usually correspond mainly to the 3' untranslated region (UTR)
XX  CC
XX  of the mRNA because they are often obtained from oligo-dT primed cDNA
XX  CC
XX  libraries. Such ESTs are not well suited for isolating cDNA sequences
XX  CC
XX  derived from the 5' ends of mRNAs and even in those cases where longer
XX  CC
XX  cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX  CC
XX  5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX  CC
XX  used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX  CC
XX  in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX  CC
XX  They are used to obtain upstream regulatory sequences and to design
XX  CC
XX  expression and secretion vectors.
XX  SQ
XX  Sequence 149 BP; 29 A; 32 C; 58 G; 29 T; 1 other;

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAC10571/rev ..
Align seg 1/1 to reverse of: AAC10571 from: 1 to: 149

1 GlyThr*****ProSer 6
|||||
75 GGAACCTCTTCTCGTCC 58

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAC1896

seq_documentation_block:
ID AAC1896 standard; cDNA; 152 BP.
XX AC
XX AAC1896;
XX DT
XX 06-OCT-2000 (first entry)

```

```

XX  Human secreted protein 5' EST, SEQ ID NO: 35971.
XX  KW
XX  Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX  KW
XX  gene therapy; chromosome mapping; ss.
XX  OS
XX  Homo sapiens.
XX  PN
XX  EPI033401-A2.
XX  PD
XX  06-SEP-2000.
XX  PF
XX  21-FEB-2000; 2000EP-0200610.
XX  PR
XX  26-FEB-1999; 99US-0122487.
XX  PA
XX  (GEST ) GENSET.
XX  PI
XX  Dumas Milne Edwards J, Duclert A, Giordano J;
XX  DR
XX  WPI; 2000-500381/45.
XX  PS
XX  Claim 1; SEQ ID 35971; 71pp + CD-ROM; English.
XX  PT
XX  The present sequence is one of a large number of 5' ESTs derived from
XX  CC
XX  mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX  CC
XX  identified within the present sequence. The 5' ESTs were prepared from
XX  CC
XX  total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX  CC
XX  sequences usually correspond mainly to the 3' untranslated region (UTR)
XX  CC
XX  of the mRNA because they are often obtained from oligo-dT primed cDNA
XX  CC
XX  libraries. Such ESTs are not well suited for isolating cDNA sequences
XX  CC
XX  derived from the 5' ends of mRNAs and even in those cases where longer
XX  CC
XX  cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX  CC
XX  5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX  CC
XX  used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX  CC
XX  in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX  CC
XX  They are used to obtain upstream regulatory sequences and to design
XX  CC
XX  expression and secretion vectors.
XX  SQ
XX  Sequence 152 BP; 34 A; 47 C; 42 G; 28 T; 1 other;

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAC31896 ..
Align seg 1/1 to: AAC31896 from: 1 to: 152

1 GlyThr*****ProSer 6
|||||
118 GGCACACGCGCMCCTCT 135

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA32011

seq_documentation_block:
ID AAA32011 standard; DNA; 153 BP.
XX AC
XX AAA32011;
XX DT
XX 05-JUL-2000 (first entry)
XX DE
XX Plant microsatellite marker #972.
XX KW
XX Plant microsatellite sequence; core repeat sequence; detection; probe;
XX KW
XX DNA polymorphism; genome mapping; physical mapping; fingerprinting;

```

```

KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Eucalyptus grandis.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX
PF 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 362; 392pp; English.
XX
CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 153 BP; 27 A; 47 C; 29 G; 48 T; 2 other;

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAA32011 ..
Align seg 1/1 to: AAA32011 from: 1 to: 153

1 GlyThr*****proSer 6
|||||
4 GGACACAGCTCCCTCG 21

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:AA10730

seq_documentation_block:
ID_ AA10730 standard; cDNA; 155 BP.
XX
AC AA10730;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 3187.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX

```

```

PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 589; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 155 BP; 43 A; 32 C; 48 G; 32 T; 0 other;

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAL10730/rev ..
Align seg 1/1 to reverse of: AAL10730 from: 1 to: 155

1 GlyThr*****proSer 6
|||||
109 GGCACCTCCTCACCACG 92

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1996.DAT:AAT39088

seq_documentation_block:
ID_ AAT39088 standard; cDNA; 171 BP.
XX
AC AAT39088;
XX
DT 21-MAY-1997 (first entry)
XX
DE Partial Coniothecium endogluconase coding sequence.
XX
KW Cellulytic enzyme; endogluconase; hydrolysis; cellulose; microorganism;
KW plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW stone-washing; cellulosic fabric; colour clarification; defibration;
KW cell wall degradation; paper pulp; debarking; fibre modification;
KW enzymatic de-inking; drainage improvement; ss.
XX
OS Coniothecium sp.
XX
PN WO9629397-A1.
XX
PD 26-SEP-1996.
XX
PF 18-MAR-1996; 96WO-DK00105.
XX
PD 12-FEB-1996; 96DK-0000137.
XX

```

```
PR 17-MAR-1995; 95DK-0000272.
PR 08-AUG-1995; 95DK-0000885.
PR 08-AUG-1995; 95DK-0000886.
PR 08-AUG-1995; 95DK-0000887.
PR 08-AUG-1995; 95DK-0000888.
XX XX (NOVO ) NOVO-NORDISK AS.
XX XX
XX XX Andersen LN, Ihara M, Kauppinen MS, Lange L, Lassen SF;
PI Nielsen RI, Schuelein M, Takagi S;
XX XX
XX XX WPI; 1996-443173/44.
XX XX P-PSDB; AAW04950.
XX XX
XX XX New endo-glucanase enzyme preparations - contg. conserved catalytic
PT regions, useful for treating fabrics, textiles, plant material or
PT paper pulp
XX XX
XX XX Example 5; Page 211-212; 316pp; English.
XX XX
XX XX AAT39076-T39100, and AAT41491 represent partial endoglucanase coding
CC sequences. The encoded peptides are examples of fragments of the
CC enzymes of the invention (see AAW04925 for an example of a full length
CC endoglucanase of the invention). The enzymes possess cellulytic
CC (particularly endoglucanase) activity. Cellulytic enzymes are involved
CC in the hydrolysis of cellulose, and are synthesised by a large number of
CC microorganisms and plants. The enzymes of the invention containing the
CC conserved catalytic regions (such as AAW04913) exhibit improved
CC performance, e.g. 50 times higher performance, compared to multiple
CC domain enzymes. The enzymes can be used for the treatment of fabrics or
CC textiles, preferably for preventing backstaining, for bio-polishing or
CC for stone-washing cellulosic fabric. They can also be used to provide
CC colour clarification for laundry. The enzymes can also be used for the
CC degradation or modification of plant material, such as cell walls. They
CC can also be used in the treatment of paper pulp preferably for
CC debarking, defibrating, fibre modification, enzymatic de-linking or
CC drainage improvement.
XX XX
XX XX Sequence 171 BP; 33 A; 67 C; 50 G; 21 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAT39088 ..
Align seg 1/1 to: AAT39088 from: 1 to: 171
1 GlyThr*****ProSer 6
|||||
26 GGAACGACACCCCTCT 43

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA74995

seq_documentation_block:
ID ABA74995 standard; DNA; 184 BP.
XX XX
XX XX ABA74995;
XX XX
XX XX 01-FEB-2002 (first entry)
XX XX
XX XX Human foetal liver single exon nucleic acid probe #23300.
XX XX
XX XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO200157277-A2.
XX XX
XX XX 09-AUG-2001.
XX XX
XX XX
```

```
XX 30-JAN-2001; 2001WO-US00669.
XX XX
XX XX 04-FEB-2000; 2000US-0180312.
XX XX 26-MAY-2000; 2000US-0207456.
XX XX 30-JUN-2000; 2000US-0608408.
XX XX 03-AUG-2000; 2000US-0632366.
XX XX 21-SEP-2000; 2000US-0234687.
XX XX 27-SEP-2000; 2000US-0236359.
XX XX 04-OCT-2000; 2000GB-0024263.
XX XX (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX WPI; 2001-483447/52.
XX XX
XX XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PT
XX XX
XX XX Claim 4; SEQ ID NO 23300; 639pp + sequence listing; English.
XX XX
XX XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
XX XX
XX XX Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fip.wipo.int/pub/published_pct_sequences.
XX XX
XX XX Sequence 184 BP; 50 A; 48 C; 41 G; 45 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x ABA74995/rev ..
Align seg 1/1 to reverse of: ABA74995 from: 1 to: 184
1 GlyThr*****ProSer 6
|||||
45 GGAACAGCTTCCCATCC 28

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA39684

seq_documentation_block:
ID ABA39684 standard; DNA; 184 BP.
XX XX
XX XX ABA39684;
XX XX
XX XX 23-JAN-2002 (first entry)
XX XX
XX XX Probe #18150 for gene expression analysis in human heart cell sample.
XX XX
XX XX Human; gene expression; heart; microarray; vascular system; probe;
XX XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX XX congenital heart disease; ss.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO200157274-A2.
XX XX
XX XX 09-AUG-2001.
XX XX
XX XX 30-JAN-2001; 2001WO-US00666.
XX XX
XX XX 04-FEB-2000; 2000US-0180312.
XX XX
```

```
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 4; SEQ ID NO 18150; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 184 BP; 50 A; 48 C; 41 G; 45 T; 0 other;
SQ

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x ABA39684/rev ..
  1 GlyThr*****ProSer 6
  |||||
  45 GGAACAGCTTCCCATCC 28

Align seg 1/1 to reverse of: ABA39684 from: 1 to: 184

seq_name: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:AAK23505
seq_documentation_block:
  ID AAK23505 standard; DNA; 184 BP.
  AC AAK23505;
  DT 05-NOV-2001 (first entry)
  DE Human brain expressed single exon probe SEQ ID NO: 23496.
  XX
  XX Human; brain expressed exon; gene expression analysis; probe;
  KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
  KW epilepsy; cancer; ss.
  XX
  XX Homo sapiens.
  OS
  XX WO200157275-A2.
  PN
  XX
  XX 09-AUG-2001.
  PD
  XX
  XX 30-JAN-2001; 2001WO-US00667.
  PF
  XX
  XX 04-FEB-2000; 2000US-0180312.
  PR
  XX 26-MAY-2000; 2000US-0207456.
  PR
```

```
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000US-0236359.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 23496; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 184 BP; 50 A; 48 C; 41 G; 45 T; 0 other;
SQ

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AAK23505/rev ..
  1 GlyThr*****ProSer 6
  |||||
  45 GGAACAGCTTCCCATCC 28

Align seg 1/1 to reverse of: AAK23505 from: 1 to: 184

seq_name: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:AAK49639
seq_documentation_block:
  ID AAK49639 standard; DNA; 184 BP.
  AC AAK49639;
  DT 06-NOV-2001 (first entry)
  DE Human bone marrow expressed single exon probe SEQ ID NO: 24196.
  XX
  XX Human; bone marrow expressed exon; gene expression analysis; probe;
  KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
  XX
  XX Homo sapiens.
  OS
  XX WO200157276-A2.
  PN
  XX
  XX 09-AUG-2001.
  PD
  XX
  XX 30-JAN-2001; 2001WO-US00668.
  PF
  XX
  XX 04-FEB-2000; 2000US-0180312.
  PR
  XX 26-MAY-2000; 2000US-0207456.
  PR
  XX 30-JUN-2000; 2000US-0608408.
  PR
  XX 03-AUG-2000; 2000US-0632366.
  PR
  XX 21-SEP-2000; 2000US-0234687.
  PR
  XX 27-SEP-2000; 2000US-0236359.
  PR
  XX 04-OCT-2000; 2000GB-0024263.
  XX
  XX (MOLE-) MOLECULAR DYNAMICS INC.
  XX PA
```

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 24196; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 184 BP; 50 A; 48 C; 41 G; 45 T; 0 other;
XX
XX alignment_scores:
XX Quality: 22.00 Length: 6
XX Ratio: 3.667 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX BASK-853-CLAIM5 x AAK49639/rev ..
XX
XX Align seg 1/1 to reverse of: AAK49639 from: 1 to: 184
XX
XX 1 GlyThr*****ProSer 6
XX |||||
XX 45 GGAACAGCTTCCCATCC 28
XX
XX seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI26748
XX
XX seq_documentation_block:
XX ID AAI26748 standard; DNA; 184 BP.
XX
XX AC AAI26748;
XX
XX DT 12-OCT-2001 (first entry)
XX
XX DE Probe #16681 for gene expression analysis in human cervical cell sample.
XX
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157278-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00670.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX PT

PS Claim 25; SEQ ID No 16681; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 184 BP; 50 A; 48 C; 41 G; 45 T; 0 other;
XX
XX alignment_scores:
XX Quality: 22.00 Length: 6
XX Ratio: 3.667 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX BASK-853-CLAIM5 x AAI26748/rev ..
XX
XX Align seg 1/1 to reverse of: AAI26748 from: 1 to: 184
XX
XX 1 GlyThr*****ProSer 6
XX |||||
XX 45 GGAACAGCTTCCCATCC 28
XX
XX seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI55522
XX
XX seq_documentation_block:
XX ID AAI55522 standard; DNA; 184 BP.
XX
XX AC AAI55522;
XX
XX DT 17-OCT-2001 (first entry)
XX
XX DE Probe #24208 used to measure gene expression in human placenta sample.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157272-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00663.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX PS Claim 25; SEQ ID No 24208; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
XX The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.

XX Sequence 184 BP; 50 A; 48 C; 41 G; 45 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAI55522/rev ..
Align seg 1/1 to reverse of: AAI55522 from: 1 to: 184

1 GlyThr*****ProSer 6
|||||
45 GGAACAGCTTCCCATCC 28

seq_name: /SDSL/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC30848

seq_documentation_block:
ID AAC30848 standard; cDNA; 189 BP.

XX AAC30848;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST; SEQ ID NO: 34923.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EF1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 9905-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 34923; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX Sequence 189 BP; 21 A; 56 C; 75 G; 34 T; 3 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAC30848/rev ..
Align seg 1/1 to reverse of: AAC30848 from: 1 to: 189

1 GlyThr*****ProSer 6
|||||
174 GGGACAGCAGCCCTTCT 157

seq_name: /SDSL/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAL06804

seq_documentation_block:
ID AAL06804 standard; DNA; 203 BP.

XX AAL06804;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 9492.

XX Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01339.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226868.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

```
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239535.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244517.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
DR
XX
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen -
PT is used in preventing, treating or ameliorating a medical condition -
XX
XX Disclosure; SEQ ID NO 9492; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
SQ Sequence 203 BP; 53 A; 43 C; 61 G; 46 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAL06804/rev ..
Align seg 1/1 to reverse of: AAL06804 from: 1 to: 203
1 GlyThr*****ProSer 6
|||||
94 GGAACACTACTGCCCCCAGT 77

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI29261
seq_documentation_block:
ID AAI29261 standard; cDNA; 203 BP.
XX
AC AAI29261;
XX
DT 12-OCT-2001 (first entry)
XX
DE Colon tumour related determined cDNA sequence for clone R0094:G04.
```

XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW Immunogenic; gene therapy; vaccine; colonic cancer; ss.
OS Homo sapiens.

XX WO200149716-A2.

XX 12-JUL-2001.

XX 29-DEC-2000; 2000WO-US35596.

XX 30-DEC-1999; 99US-0476296.

XX 10-JAN-2000; 2000US-0480321.

XX 13-FEB-2000; 2000US-0504629.

XX 06-MAR-2000; 2000US-0519444.

XX 19-MAY-2000; 2000US-0575251.

XX 29-JUN-2000; 2000US-0609448.

XX 28-AUG-2000; 2000US-0649811.

XX (CORI-) CORIXA CORP.

XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;

XX King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.

XX Colon tumor associated proteins and nucleic acids useful for the

XX prevention, diagnosis and treatment of colonic cancer -

XX Claim 2; Page 362; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AAI28460 to AAI29512
CC and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.

XX Sequence 203 BP; 46 A; 65 C; 58 G; 34 T; 0 other;

alignment_scores: Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AAI29261/rev

Align seg 1/1 to reverse of: AAI29261 from: 1 to: 203

1 GlyThr*****ProSer 6

|||||

122 GGGACAGGCTCCATCA 105

seq_name: /STDsl/gcgdata/geneseq/geneseq-enbl/NA2001A.DAT:AAI62685

seq_documentation_block:

ID AAI62685 standard; DNA: 203 BP.

XX AAI62685;

XX 19-OCT-2001 (first entry)

XX Human breast or ovarian antigen genomic DNA SEQ ID NO: 335.

XX Human; breast antigen; ovarian antigen; cancer; metastasis; gene therapy;

XX Homo sapiens.

XX WO200155324-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01344.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 14-AUG-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226868.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

XX 30-AUG-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

XX 01-SEP-2000; 2000US-0229344.

XX 01-SEP-2000; 2000US-0229345.

XX 03-SEP-2000; 2000US-0229509.

XX 05-SEP-2000; 2000US-0229513.

XX 06-SEP-2000; 2000US-0230437.

XX 08-SEP-2000; 2000US-0230438.

XX 08-SEP-2000; 2000US-0231242.

XX 08-SEP-2000; 2000US-0231243.

XX 08-SEP-2000; 2000US-0231244.

XX 08-SEP-2000; 2000US-0231413.

XX 08-SEP-2000; 2000US-0231414.

XX 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 03-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-488785/53.
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
XX treating and/or preventing human diseases and disorders -
XX Disclosure; SEQ ID NO: 335; 520pp + Sequence Listing; English.
XX The present invention provides the protein and coding sequences of a
XX number of ovarian and breast antigens. These are shown in
XX AA162467-AA162572 and AAM42240-AAM42345. The sequences can be used in the
XX diagnosis, prevention and treatment of breast and ovarian cancers, and
XX their metastases. The present sequence is a genomic sequence of the
XX invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 203 BP; 53 A; 43 C; 61 G; 46 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAI62685/rev ..
Align seg 1/1 to reverse of: AAI62685 from: 1 to: 203
1 GlyThr*****ProSer 6
|||||
94 GGAACACTACTGCCCCCAGT 77

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK65547
seq_documentation_block:
ID AAK65547 standard; DNA; 218 BP.
XX AC AAK65547;
XX DT 06-NOV-2001 (first entry)
XX DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:20359.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX WO200157182-A2.
XX PN 09-AUG-2001.
XX PD 17-JAN-2001; 2001WO-US01354.
XX PF

PT metastasis -
XX Disclosure; SEQ ID NO 20359; 3071pp + Sequence Listing; English.
PS
CC
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK34942 to AAK34950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 218 BP; 36 A; 63 C; 93 G; 26 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAK65547/rev ..

Align seg 1/1 to reverse of: AAK65547 from: 1 to: 218

1 GlyThr*****ProSer 6
|||||||
87 GGCACCGCAGCTCCGAGT 70

seq_name: /SDSL/gcgdata/geneseq/geneseq-embl/NA2000.DAT:AAK01664

seq_documentation_block:

ID AAK01664 standard; cDNA; 221 BP.

AC AAK01664;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 1662.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

PN EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR P-PSDB; AAG01658.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 1662; 71pp + CD-ROM; English.
PS

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

XX SQ Sequence 221 BP; 55 A; 46 C; 70 G; 50 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAK01664/rev ..

Align seg 1/1 to reverse of: AAK01664 from: 1 to: 221

1 GlyThr*****ProSer 6
|||||||
90 GGTACTGCAGCCCTTCA 73

seq_name: /SDSL/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABL25507

seq_documentation_block:

ID ABL25507 standard; DNA; 225 BP.

AC ABL25507;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 27994.

DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEXE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PS Claim 1; SEQ ID NO 27994; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 225 BP; 48 A; 62 C; 61 G; 54 T; 0 other;

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x ABL25507/rev ..

Align seg 1/1 to reverse of: ABL25507 from: 1 to: 225

1 GlyThr*****ProSer 6
|||||||
109 GGAAGTCTCTCTCTCTCT 92

seq_name: /SDSI/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:AAS34057

seq_documentation_block:

ID AAS34057 standard; cDNA; 233 Bp.

XX AAS34057;

AC AAS34057;

XX 17-DEC-2001 (first entry)

XX Human cDNA encoding a novel foetal antigen, SEQ ID No 581.

XX Human; foetal tissue antigen; ss; antiinflammatory; neuroprotective;
KW immunomodulator; cardiovascular; cytostatic; nephrothropic;
KW cardiovascular; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; breast neoplasm; cancer;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; food additive.

XX Homo sapiens.

XX WO200155312-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01321.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0246417.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488782/53.
DR P-PSDB; AAU21237.
XX
PT New polynucleotides and polypeptides for diagnosing, treating,
PT preventing or prognosing e.g. diseases or disorders of the nervous,
PT musculoskeletal, excretory, gastrointestinal, reproductive, and
PT respiratory systems -
XX
XX Claim 1; SEQ ID No 581; 642pp; English.
PS
XX The invention relates to novel nucleic acids encoding novel human foetal
CC antigens. The nucleic acids and proteins are used to prevent, treat (e.g.
CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. The antibodies to the antigens can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders

CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. Numerous
CC examples of diseases and disorders treated by the nucleic acids and
CC proteins are given in the specification. The present sequence

alignment_scores: Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AAS34057/rev ..

Align seg 1/1 to reverse of: AAS34057 from: 1 to: 233

1 GlyThr*****ProSer 6
|||||||
62 GGACATCTTCCCTCT 45

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.AAT20850

seq_documentation_block:

ID AAT20850 standard; cDNA to mRNA; 236 BP.

XX

AC AAT20850;

XX

DT 10-JUL-1996 (first entry)

XX

DE Human gene signature HUMGS02100.

XX

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX WO9514772-A1.

XX 01-JUN-1995.

XX 11-NOV-1994; 94WO-JP01916.

XX 12-NOV-1993; 93JP-0355504.

XX (MATS/) MATSUBARA K.

XX (OKUB/) OKUBO K.

XX Matsubara K, Okubo K;

XX WPI; 1995-206931/27.

XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX for diagnosis of abnormal cell function, by preparing cDNA that
XX reflects relative abundance of corresp. mRNA in specific human
XX tissues

PS Claim 1; Page 741; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

XX
SQ Sequence 236 BP; 46 A; 79 C; 46 G; 56 T; 9 other;

alignment_scores:

Quality:	22.00	Length:	6
Ratio:	3.667	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

BASK-853-CLAIM5 x AAT20850 ..

Align seg 1/1 to: AAT20850 from: 1 to: 236

1 GlyThr*****ProSer 6
|||||
30 GGGACTACCTCTCTTCA 47

OM of: BASK-853-CLAIM5 to: Issued_Patents_NA:* out_format : pfs

Date: Sep 24, 2002 2:13 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=framer_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/BASKR853075/runat_17092002_103408_1838/app_query.fasta_1.98
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=mod.rni
-GAPOP=12,000 -GAPEXT=4,000 -MINMATCH=0,100 -LOOPCL=0,000
-LOOPEXT=0,000 -QGAPOP=4,500 -QGAPEXT=0,050 -XGAPOP=10,000
-XGAPEXT=0,500 -DELOP=6,000 -FGAPEXT=7,000 -YGAPOP=10,000
-MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=50 -DOALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=BASKR853075 @cgn1_1.72 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_MLUPXY -WAIT -THREADS=1

Search information block:

Query: BASK-853-CLAIM5
Query length: 6
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 82.560000

score_list:

Sequence	Strd Orig	Zscore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-143-634-11 +	22.00	109.58	204.12	58	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-651-136C-51 +	22.00	101.31	589.69	171	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-204-740-5 +	22.00	97.73	933.19	273	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-167A-5 +	22.00	97.73	933.19	273	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-395-5 +	22.00	97.73	933.19	273	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-172-711-19 +	22.00	97.73	933.19	273	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-416-833-5 +	22.00	97.73	933.19	273	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-143-634-11 +	22.00	97.73	933.19	273	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-946-914-29 +	22.00	97.70	936.54	274	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-716-942-10 -	22.00	97.32	983.47	288	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-199-637A-74 +	22.00	97.01	1.0e+03	300	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-199-637A-76 -	22.00	96.85	1.0e+03	306	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-420-235B-36 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-343-101A-2 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-747-887-3 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-708-678B-3 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-728-322A-3 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-183-688-2 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-624-36 +	22.00	96.28	1.1e+03	330	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-480-173A-36 +	22.00	95.48	1.2e+03	366	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-484-408A-55 +	22.00	95.48	1.2e+03	366	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-480-173A-53 +	22.00	95.30	1.3e+03	375	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-484-408A-53 +	22.00	95.30	1.3e+03	375	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-179-72 +	22.00	94.38	1.4e+03	423	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-879-685B-3 -	22.00	93.27	1.7e+03	489	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-579A-1 +	22.00	93.10	1.7e+03	500	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-601-184-1 +	22.00	93.10	1.7e+03	500	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-073-297-15 +	22.00	92.93	1.7e+03	511	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-000-630C-26 -	22.00	92.59	1.8e+03	534	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-862-730C-26 -	22.00	92.59	1.8e+03	534	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-923-454A-22 -	22.00	91.87	2.0e+03	587	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-073-297-25 +	22.00	91.73	2.0e+03	598	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-1142 +	22.00	90.79	2.3e+03	676	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-348 +	22.00	90.60	2.3e+03	693	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-1005 +	22.00	90.49	2.4e+03	703	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-613-942-1 +	22.00	90.38	2.4e+03	713	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-673 -	22.00	90.33	2.4e+03	718	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-020-956-37 -	22.00	89.89	2.5e+03	760	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-607-37 -	22.00	89.89	2.5e+03	760	!

/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-439-313-37 - 22.00 89.89 2.5e+03 760
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-154-083-23 - 22.00 89.58 2.7e+03 792
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-453 + 22.00 89.56 2.7e+03 794
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-199-637A-30 + 22.00 89.52 2.7e+03 798
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-051-935A-49 + 22.00 89.35 2.7e+03 816
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-020-956-14 - 22.00 89.35 2.7e+03 816
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-607-14 - 22.00 89.35 2.7e+03 816
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-439-313-14 - 22.00 89.35 2.7e+03 816
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-361-707-104 + 22.00 89.30 2.7e+03 821
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-434-774-5 - 22.00 89.02 2.9e+03 852

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-143-634-11

seq_documentation_block:

; Sequence 11, Application US/09143634
; Patent No. 6214602
; GENERAL INFORMATION:
; APPLICANT: Zdanovsky, Alexey G.
; TITLE OF INVENTION: EXPRESSION OF CLOSTRIDIAL TOXINS AND PROTEINS
; FILE REFERENCE: PRMG-03193
; CURRENT APPLICATION NUMBER: US/09/143,634
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 58
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-143-634-11

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x US-09-143-634-11 ..

Align seg 1/1 to: US-09-143-634-11 from: 1 to: 58

1 Glythr*****Proser 6

|||||||
5 GGGACATCATCACCATCA 22

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-651-136C-51

seq_documentation_block:

; Sequence 51, Application US/08651136C
; Patent No. 6001639
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; APPLICANT: Andersen, Lene N.
; APPLICANT: Lassen, Soren F.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Lange, Lene
; APPLICANT: Nielsen, Ruby I.
; APPLICANT: Ihara, Michiko
; APPLICANT: Takagi, Shinobu
; TITLE OF INVENTION: No. 6001639el Endoglucanases
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136C
FILING DATE: 21-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366,200-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..171
US-08-651-136C-51

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-651-136C-51 ..

Align seg 1/1 to: US-08-651-136C-51 from: 1 to: 171

1 GlyThr*****ProSer 6
|||||
26 GGAACGACACCCCTCT 43

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-204-740-5

seq_documentation_block:
Sequence 5, Application US/08204740
Patent No. 5753432
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-C
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-204-740-5

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-204-740-5 ..

Align seg 1/1 to: US-08-204-740-5 from: 1 to: 273

1 GlyThr*****ProSer 6
|||||
202 GGCACAGCGAGCCCTCA 219

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-167A-5

seq_documentation_block:
Sequence 5, Application US/09081167A
Patent No. 6083745

GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083745nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```
; MOLECULE TYPE: cdna
US-09-081-167A-5

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-081-167A-5
Align seg 1/1 to: US-09-081-167A-5 from: 1 to: 273

  1 GlyThr*****ProSer 6
  |||||
  202 GGCACAGCGAGCCCTCA 219

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-395-5

seq_documentation_block:
; Sequence 5, Application US/09081395
; Patent No. 6083746
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,395
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083746nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-081-395-5

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-081-395-5
Align seg 1/1 to: US-09-081-395-5 from: 1 to: 273

  1 GlyThr*****ProSer 6
  |||||
  202 GGCACAGCGAGCCCTCA 219

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-395-5

seq_documentation_block:
; Sequence 5, Application US/09081395
; Patent No. 6160105
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Zweigler, Gary B.
; APPLICANT: Panzer, Scott R.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: MONITORING TOXICOLOGICAL RESPONSES
; FILE REFERENCE: PA-0011 US
; CURRENT APPLICATION NUMBER: US/09/172,711
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PERL Program
; SEQ ID NO 19
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 52, 139, 249
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE:
; OTHER INFORMATION: 700303390H1
US-09-172-711-19

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-172-711-19
Align seg 1/1 to: US-09-172-711-19 from: 1 to: 273

  1 GlyThr*****ProSer 6
  |||||
  221 GGGACTTCCAGTCCATCC 238

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-416-833-5

seq_documentation_block:
; Sequence 5, Application US/09416833
; Patent No. 6197521
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416,833
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/204,740
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6197521nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-416-833-5

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-416-833-5 ..
Align seg 1/1 to: US-09-416-833-5 from: 1 to: 273

1 GlyThr*****ProSer 6
|||||
202 GGGACAGCGAGCCTTCA 219

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-02521-5

seq_documentation_block:
; Sequence 5, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; PCT-US95-02521-5

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
```

```
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x PCT-US95-02521-5 ..
Align seg 1/1 to: PCT-US95-02521-5 from: 1 to: 273

1 GlyThr*****ProSer 6
|||||
202 GGGACAGCGAGCCTTCA 219

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-946-914-29

seq_documentation_block:
; Sequence 29, Application US/08946914
; Patent No. 6027916
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,914
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,093
; FILING DATE: 09-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0560001/EKS/SGW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-946-914-29

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-946-914-29 ..
Align seg 1/1 to: US-08-946-914-29 from: 1 to: 274

1 GlyThr*****ProSer 6
|||||
186 GGAACACATCGCCTCC 203

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-716-942-10
```

```
seq_documentation_block:
; Sequence 10, Application US/08716942
; Patent No. 5849491
; GENERAL INFORMATION:
; APPLICANT: Terragen Diversity, Inc.
; APPLICANT: Radomski, Christopher C. A.
; APPLICANT: Seow, Kah Tong
; APPLICANT: Warren, R. Antony J.
; APPLICANT: Yap, Wai Ho
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: METHOD FOR ISOLATING XYLANASE GENE
; TITLE OF INVENTION: SEQUENCES FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD AND
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppedahl & Larson
; STREET: 1992 Commerce Street, Suite 309
; CITY: Yorktown Heights
; STATE: NY
; COUNTRY: USA
; ZIP: 10598-4412
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS 5.0.
; SOFTWARE: Word Perfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,942
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/004,157
; FILING DATE: 20-Sep-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Marina T. Larson
; REGISTRATION NUMBER: 32,038
; REFERENCE/DOCKET NUMBER: TERR-P-001
; TELEPHONE: (914) 245-3252
; TELEFAX: (914) 962-4330
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288
; TYPE: nucleic acid
; STRANDEDNESS: DOUBLE
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: fragment of xylanase gene from degenerate primer
; NAME/KEY: amplification of soil DNA
US-08-716-942-10

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-716-942-10 reverse of: US-08-716-942-10 from: 1 to: 288
..
1 GlyThr*****ProSer 6
|||||
274 GGGACAGCAGCAGCCGTCG 257

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-199-637A-74

seq_documentation_block:
; Sequence 76, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-74

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-199-637A-74 ..
Align seg 1/1 to: US-09-199-637A-74 from: 1 to: 300
..
1 GlyThr*****ProSer 6
|||||
123 GGAACGACATCACCATCG 140

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-199-637A-76

seq_documentation_block:
; Sequence 76, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-76
```

```

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x US-09-199-637A-76/rev ..
  Align seg 1/1 to reverse of: US-09-199-637A-76 from: 1 to: 306

      1 GlyThr*****ProSer 6
      |||||
152 GGGACGACATCACCATCG 135

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-420-235B-36

seq_documentation_block:
; Sequence 36, Application US/08420235B
; Patent No. 5801042
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/420,235B
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
US-08-420-235B-36

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x US-08-420-235B-36 ..
  Align seg 1/1 to: US-08-420-235B-36 from: 1 to: 330

      1 GlyThr*****ProSer 6
      |||||

```

```

54 GGGACAGCAACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-343-101A-2

seq_documentation_block:
; Sequence 2, Application US/08343101A
; Patent No. 5830759
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
; TITLE OF INVENTION: Virus Sequences And Uses Thereof
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,101A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
US-08-343-101A-2

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x US-08-343-101A-2 ..
  Align seg 1/1 to: US-08-343-101A-2 from: 1 to: 330

      1 GlyThr*****ProSer 6
      |||||
54 GGGACAGCAACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-747-887-3

seq_documentation_block:
; Sequence 3, Application US/08747887
; Patent No. 5853734
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.

```



```
; TITLE OF INVENTION: Glycoprotein L And Glycoprotein M From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA Encoding
; TITLE OF INVENTION: Same And Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/747,887
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52429
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-08-747-887-3

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-747-887-3
..
Align seg 1/1 to: US-08-747-887-3 from: 1 to: 330

1 GlyThr*****ProSer 6
|||||
54 GGGACAGCACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/5b_COMB.seq:US-08-708-678B-3

seq_documentation_block:
; Sequence 3, Application US/08708678B
; Patent No. 5859225
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Virion protein 26 From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
```

```
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,678B
; FILING DATE: September 5, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51945/JPW/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-08-708-678B-3

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-708-678B-3
..
Align seg 1/1 to: US-08-708-678B-3 from: 1 to: 330

1 GlyThr*****ProSer 6
|||||
54 GGGACAGCACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/5b_COMB.seq:US-08-728-323A-3

seq_documentation_block:
; Sequence 3, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,323A
```

```
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-08-728-323A-3

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-728-323A-3 ..
Align seg 1/1 to: US-08-728-323A-3 from: 1 to: 330

1 GlyThr*****ProSer 6
|||||
54 GGGACAGCAACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-183-688-2

seq_documentation_block:
; Sequence 2, Application US/09183688
; Patent No. 6093550
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
; TITLE OF INVENTION: Virus Sequences And Uses Thereof
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/183,688
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/343,101
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526
; INFORMATION FOR SEQ ID NO: 2:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-09-183-688-2

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-183-688-2 ..
Align seg 1/1 to: US-09-183-688-2 from: 1 to: 330

1 GlyThr*****ProSer 6
|||||
54 GGGACAGCAACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-624-36

seq_documentation_block:
; Sequence 36, Application US/08793624C
; Patent No. 6150093
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma Virus Sequences And
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 45185-C-PCT-US/JPW
; CURRENT APPLICATION NUMBER: US/08/793,624C
; CURRENT FILING DATE: 1997-02-18
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 36
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
; US-08-793-624-36

alignment_scores:
  Quality: 22.00 Length: 6
  Ratio: 3.667 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-793-624-36 ..
Align seg 1/1 to: US-08-793-624-36 from: 1 to: 330

1 GlyThr*****ProSer 6
|||||
54 GGGACAGCAACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-10194-36

seq_documentation_block:
; Sequence 36, Application PC/TUS9510194
; GENERAL INFORMATION:
; APPLICANT: The trustees of Columbia University in the City of New York
; APPLICANT: City
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
```

```

; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10194
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPW/MS-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; PCT-US95-10194-36

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x PCT-US95-10194-36  ..
  Align seg 1/1 to: PCT-US95-10194-36 from: 1 to: 330

  1 GlyThr*****ProSer 6
  |||||
  54 GGCACAGCACACCCAGC 71

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-480-173A-55

seq_documentation_block:
; Sequence 55, Application US/08480173A
; Patent No. 6072049
; GENERAL INFORMATION:
; APPLICANT: Thoma, Hans A
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Popovich & Wiles, P.A.
; STREET: 80 S. 8th Street, Suite 1902
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,173A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Popovich, Thomas E
; REGISTRATION NUMBER: 30,099
; REFERENCE/DOCKET NUMBER: MED1003USD4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-334-8991
; TELEFAX: 612-334-8994
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..3
; OTHER INFORMATION: /note= "Pre-S1 start codon"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 358..360
; OTHER INFORMATION: /note= "Pre-S2 start codon"
; NAME/KEY: CDS
; LOCATION: 1..366
; US-08-480-173A-55

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-480-173A-55  ..
  Align seg 1/1 to: US-08-480-173A-55 from: 1 to: 366

  1 GlyThr*****ProSer 6
  |||||
  128 GGGACTTCAACCCCATCA 145

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-484-408A-55

seq_documentation_block:
; Sequence 55, Application US/08484408A
; Patent No. 6117653
; GENERAL INFORMATION:
; APPLICANT: Thoma, Hans A
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Popovich & Wiles, P.A.
; STREET: 80 S. 8th Street, Suite 1902
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,408A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Popovich, Thomas E
; REGISTRATION NUMBER: 30,099
; REFERENCE/DOCKET NUMBER: MED1003USD4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-334-8991
; TELEFAX: 612-334-8994

```

```

; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..3
; OTHER INFORMATION: /note= "Pre-S1 start codon"
;
; NAME/KEY: misc_feature
; LOCATION: 358..360
; OTHER INFORMATION: /note= "Pre-S2 start codon"
;
; NAME/KEY: CDS
; LOCATION: 1..366
; US-08-484-408A-55

```

alignment_scores:

Quality:	22.00	Length:	6
Ratio:	3.667	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

BASK-853-CLAIM5 x US-08-484-408A-55 ..

Align seg 1/1 to: US-08-484-408A-55 from: 1 to: 366

```

1 GlyThr*****ProSer 6
|||||
128 GGGACTTCAACCCATCA 145

```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-480-173A-53

seq_documentation_block:

```

; Sequence 53, Application US/08480173A
; Patent No. 6072049
; GENERAL INFORMATION:
; APPLICANT: Thoma, Hans A
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Popovich & Wiles, P.A.
; STREET: 80 S. 8th Street, Suite 1902
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,173A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Popovich, Thomas E
; REGISTRATION NUMBER: 30,099
; REFERENCE/DOCKET NUMBER: MED1003USD4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-334-8994
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..3
; OTHER INFORMATION: /note= "Pre-S2 start codon"
;
; NAME/KEY: misc_feature
; LOCATION: 358..360
; OTHER INFORMATION: /note= "Pre-S2 start codon"
;
; NAME/KEY: CDS
; LOCATION: 1..375
; US-08-480-173A-53

```

alignment_scores:

Quality:	22.00	Length:	6
Ratio:	3.667	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

BASK-853-CLAIM5 x US-08-480-173A-53 ..

Align seg 1/1 to: US-08-480-173A-53 from: 1 to: 375

```

1 GlyThr*****ProSer 6
|||||
128 GGGACTTCAACCCGTC A 145

```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-484-408A-53

seq_documentation_block:

```

; Sequence 53, Application US/08484408A
; Patent No. 6117653
; GENERAL INFORMATION:
; APPLICANT: Thoma, Hans A
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Popovich & Wiles, P.A.
; STREET: 80 S. 8th Street, Suite 1902
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,408A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Popovich, Thomas E
; REGISTRATION NUMBER: 30,099
; REFERENCE/DOCKET NUMBER: MED1003USD4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-334-8994
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..3
; OTHER INFORMATION: /note= "Pre-S2 start codon"
;
; FEATURE:

```

```

; NAME/KEY: misc.feature
; LOCATION: 358..360
; OTHER INFORMATION: /note= "Pre-S2 start codon"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..375
; US-08-484-408A-53

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-484-408A-53 ..
Align seg 1/1 to: US-08-484-408A-53 from: 1 to: 375

1 GlyThr*****ProSer 6
|||||
128 GGGACTTCAACCCGTCA 145

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-179-72

seq_documentation_block:
; Sequence 72, Application US/08470179
; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D, Wai Mun
; TITLE OF INVENTION: Method and Compositions for
; TITLE OF INVENTION: Identification of Species in a Sample
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask, Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "gyrA gene segment"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Leptonema illini
; US-08-470-179-72

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-470-179-72 ..
Align seg 1/1 to: US-08-470-179-72 from: 1 to: 423

1 GlyThr*****ProSer 6
|||||
59 GGCACGACGCCCTTCG 76

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-879-685B-3

seq_documentation_block:
; Sequence 3, Application US/07879685B
; Patent No. 5296383
; GENERAL INFORMATION:
; APPLICANT: DAIKIN INDUSTRIES, LTD.
; TITLE OF INVENTION: A human centromere antigen
; TITLE OF INVENTION: polypeptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Umeda Center Building, 4-12
; STREET: Nakazaki-nishi, 2-chome
; CITY: Kita-ku
; STATE: Osaka
; COUNTRY: Japan
; ZIP: 530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/879,685B
; FILING DATE: 19920507
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-102517
; FILING DATE: 08-May-1991
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 to 489
; IDENTIFICATION METHOD: by experiment
; US-07-879-685B-3

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-07-879-685B-3/rev ..
Align seg 1/1 to reverse of: US-07-879-685B-3 from: 1 to: 489

1 GlyThr*****ProSer 6
|||||
308 GGCACCTCATCACCATCC 291

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-037-579A-1
```

```

Ratio: 3.667      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-470-179-72 ..
Align seg 1/1 to: US-08-470-179-72 from: 1 to: 423

1 GlyThr*****ProSer 6
|||||
59 GGCACGACGCCCTTCG 76

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-879-685B-3

seq_documentation_block:
; Sequence 3, Application US/07879685B
; Patent No. 5296383
; GENERAL INFORMATION:
; APPLICANT: DAIKIN INDUSTRIES, LTD.
; TITLE OF INVENTION: A human centromere antigen
; TITLE OF INVENTION: polypeptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Umeda Center Building, 4-12
; STREET: Nakazaki-nishi, 2-chome
; CITY: Kita-ku
; STATE: Osaka
; COUNTRY: Japan
; ZIP: 530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/879,685B
; FILING DATE: 19920507
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-102517
; FILING DATE: 08-May-1991
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 to 489
; IDENTIFICATION METHOD: by experiment
; US-07-879-685B-3

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-07-879-685B-3/rev ..
Align seg 1/1 to reverse of: US-07-879-685B-3 from: 1 to: 489

1 GlyThr*****ProSer 6
|||||
308 GGCACCTCATCACCATCC 291

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-037-579A-1
```

```
seq_documentation_block:
; Sequence 1, Application US/08037579A
; Patent No. 552537
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Max, Edward E
; APPLICANT: Saxon, Andrew
; TITLE OF INVENTION: Ige ISOFORMS AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/037,579A
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-57950/BIR UCLA-233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..329
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 359..439
; US-08-037-579A-1

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-037-579A-1 ..
Align seg 1/1 to: US-08-037-579A-1 from: 1 to: 500
1 GlyThr*****ProSer 6
|||||
58 GGGACAAAGCGCACCTCG 75

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-601-184-1

seq_documentation_block:
; Sequence 1, Application US/08601184
; Patent No. 6043345
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Max, Edward E
; APPLICANT: Saxon, Andrew
; TITLE OF INVENTION: Ige ISOFORMS AND METHODS OF USE
```

```
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MSDOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601,184
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A-57950-1/PJS UCLA233-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..329
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 359..439
; US-08-601-184-1

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-601-184-1 ..
Align seg 1/1 to: US-08-601-184-1 from: 1 to: 500
1 GlyThr*****ProSer 6
|||||
58 GGGACAAAGCGCACCTCG 75

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-073-297-15

seq_documentation_block:
; Sequence 15, Application US/09073297
; Patent No. 6255114
; GENERAL INFORMATION:
; APPLICANT: LIGHTNER, JONATHAN EDWARD
; TITLE OF INVENTION: STARCH BIOSYNTHETIC ENZYMES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
```

```
; ; COMPUTER: IBM PC COMPATIBLE
; ; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; ; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/09/073,297
; ; FILING DATE:
; ; CLASSIFICATION:
; ; PRIOR APPLICATION DATA:
; ; APPLICATION NUMBER: 08/852615
; ; FILING DATE: MAY 7, 1997
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: MAJARIAN, WILLIAM R.
; ; REGISTRATION NUMBER: 41,173
; ; REFERENCE/DOCKET NUMBER: BB1083-B
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: 302-992-4926
; ; TELEFAX: 302-773-0164
; ; INFORMATION FOR SEQ ID NO: 15:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 511 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: cDNA
; ; IMMEDIATE SOURCE:
; ; CLONE: rrl.pk0070.e9
; ; US-09-073-297-15

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-073-297-15 ..
Align seg 1/1 to: US-09-073-297-15 from: 1 to: 511

1 GlyThr*****ProSer 6
|||||
363 GGCACCGCATCCCAAGC 380

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-000-630C-26

seq_documentation_block:
; Sequence 26, Application US/09000630C
; Patent No. 6018029
; GENERAL INFORMATION:
; APPLICANT: Fuller, Gerald M
; APPLICANT: Fuentes, Nelson L.
; TITLE OF INVENTION: DNA Encoding Canine Interleukin-1 Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Douglas C Murdock/ Bradley, Arant, Rose & White
; STREET: 2001 Park Place, Suite 1400
; CITY: Birmingham
; STATE: Alabama
; COUNTRY: USA
; ZIP: 35203-2736
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch,
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,630C
; FILING DATE: 5/23/97
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: rabbit IL-1ra DNA sequence
; US-08-862-730C-26

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-862-730C-26/rev ..
Align seg 1/1 to reverse of: US-08-862-730C-26 from: 1 to: 534
```

```
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 534 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: rabbit IL-1ra DNA sequence
; ; US-09-000-630C-26

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-000-630C-26/rev ..
Align seg 1/1 to reverse of: US-09-000-630C-26 from: 1 to: 534

1 GlyThr*****ProSer 6
|||||
274 GGCACACTTCCCCCTCT 257

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-862-730C-26

seq_documentation_block:
; Sequence 26, Application US/08862730C
; Patent No. 6063600
; GENERAL INFORMATION:
; APPLICANT: Fuller, Gerald M
; APPLICANT: Fuentes, Nelson L.
; TITLE OF INVENTION: DNA Encoding Canine Interleukin-1 Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Douglas C Murdock/ Bradley, Arant, Rose & White
; STREET: 2001 Park Place, Suite 1400
; CITY: Birmingham
; STATE: Alabama
; COUNTRY: USA
; ZIP: 35203-2736
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch,
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,730C
; FILING DATE: 5/23/97
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: rabbit IL-1ra DNA sequence
; US-08-862-730C-26

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-862-730C-26/rev ..
Align seg 1/1 to reverse of: US-08-862-730C-26 from: 1 to: 534

1 GlyThr*****ProSer 6
|||||
274 GGCACACTTCCCCCTCT 257
```

```
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-923-454A-22
seq_documentation_block:
; Sequence 22, Application US/08923454A
; Patent No. 6004794
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; APPLICANT: Livi, George
; APPLICANT: Karian, Eric
; APPLICANT: Clinkenbeard, Helen
; APPLICANT: Browne, Michael
; APPLICANT: Southan, Christopher
; TITLE OF INVENTION: HUMAN SERINE PROTEASE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,454A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/025436
; FILING DATE: 06-SEPT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-923-454A-22

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-923-454A-22/rev ..
Align seg 1/1 to reverse of: US-08-923-454A-22 from: 1 to: 587

1 GlyThr*****ProSer 6
|||||
374 GGGACCGCTAGCCCTTC 357

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-073-297-25
seq_documentation_block:
; Sequence 25, Application US/09073297
```

```
; Patent No. 6255114
; GENERAL INFORMATION:
; APPLICANT: LIGHTNER, JONATHAN EDWARD
; TITLE OF INVENTION: STARCH BIOSYNTHETIC ENZYMES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,297
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/852615
; FILING DATE: MAY 7, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MAJARIAN, WILLIAM R.
; REGISTRATION NUMBER: 41,173
; REFERENCE/DOCKET NUMBER: BB-1083-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4926
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: wleln.pk0056.b2
; US-09-073-297-25

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-073-297-25 ..
Align seg 1/1 to: US-09-073-297-25 from: 1 to: 598

1 GlyThr*****ProSer 6
|||||
192 GGCACCGCATCCCAACG 209

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-1142
seq_documentation_block:
; Sequence 1142, Application US/08998416
; Patent No. 6239284
; GENERAL INFORMATION:
; APPLICANT: Philippssen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYVPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
```



```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 676 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: PAG1694UP
; US-08-998-416-1142

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-998-416-1142 ..
1 GlyThr*****ProSer 6
|||||
359 GGGACAACGACCCCTCA 376

Align seg 1/1 to: US-08-998-416-1142 from: 1 to: 676

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-348

seq_documentation_block:
; Sequence 348, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
```

```
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(693)
; OTHER INFORMATION: n = A,T,C or G
; US-09-328-111-348

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-328-111-348 ..
1 GlyThr*****ProSer 6
|||||
227 GGGACATCTGCGCCCTCC 244

Align seg 1/1 to: US-09-328-111-348 from: 1 to: 693

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-1005

seq_documentation_block:
; Sequence 1005, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtile, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPITII
; AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
```

```

; INFORMATION FOR SEQ ID NO: 1005:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1617RP
US-08-998-416-1005

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-998-416-1005 ..
Align seg 1/1 to: US-08-998-416-1005 from: 1 to: 703

1 GlyThr*****ProSer 6
|||||
98 GGCACATCTGCGCCAGT 115

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-613-942-1

seq_documentation_block:
; Sequence 1, Application US/08613942
; Patent No. 5866791
; GENERAL INFORMATION:
; APPLICANT: HOLT, KAREN A.
; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/613,942
; FILING DATE: 13-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9521848.3
; FILING DATE: 25-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222340/SEE 45024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-613-942-1

; INFORMATION FOR SEQ ID NO: 1005:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1617RP
US-08-998-416-1005

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-613-942-1 ..
Align seg 1/1 to: US-08-613-942-1 from: 1 to: 713

1 GlyThr*****ProSer 6
|||||
289 GGTACGTCTGCGCGAGC 306

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-673

seq_documentation_block:
; Sequence 673, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtie, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPYII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NO. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 673:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 718 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1448RP
US-08-998-416-673

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
```

```
alignment_block:
BASK-853-CLAIM5 x US-08-998-416-673/rev ..
Align seg 1/1 to reverse of: US-08-998-416-673 from: 1 to: 718
1 GlyThr*****ProSer 6
470 GGCACAAGTCCCATCA 453

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-020-956-37

seq_documentation_block:
; Sequence 37, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 37:
; LENGTH: 760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-020-956-37

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-020-956-37/rev ..
Align seg 1/1 to reverse of: US-09-020-956-37 from: 1 to: 760
1 GlyThr*****ProSer 6
91 GGTACTACACCCCTTCA 74

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-607-37

seq_documentation_block:
; Sequence 37, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
```

```
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-030-607-37

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-030-607-37/rev ..
Align seg 1/1 to reverse of: US-09-030-607-37 from: 1 to: 760
1 GlyThr*****ProSer 6
91 GGTACTACACCCCTTCA 74

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-439-313-37

seq_documentation_block:
; Sequence 37, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
```

```

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(760)
; OTHER INFORMATION: n = A,T,C or G
US-09-439-313-37

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-439-313-37/rev ..

Align seg 1/1 to reverse of: US-09-439-313-37 from: 1 to: 760

1 GlyThr*****ProSer 6
|||||
91 GGTACTACAAACCCCTCA 74

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-154-083-23

seq_documentation_block:
; Sequence 23, Application US/09154083
; Patent No. 6150513
; GENERAL INFORMATION:
; APPLICANT: Wu, Kai
; TITLE OF INVENTION: Polyketide Synthase Enzymes and Recombinant DNA
; FILE REFERENCE: 30062-20022.00 Polyketide Synthase Enz
; CURRENT APPLICATION NUMBER: US/09/154,083
; CURRENT FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
US-09-154-083-23

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-154-083-23/rev ..

Align seg 1/1 to reverse of: US-09-154-083-23 from: 1 to: 792

1 GlyThr*****ProSer 6
|||||
145 GGTACAGCAGCCCCAGC 128

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-453

seq_documentation_block:
; Sequence 453, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippssen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp

; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalls Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/GC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 453:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 794 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1327UP
US-08-998-416-453

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-998-416-453 ..

Align seg 1/1 to: US-08-998-416-453 from: 1 to: 794

1 GlyThr*****ProSer 6
|||||
138 GGGACAAACAGCACCAGC 155

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-199-637A-30

seq_documentation_block:
; Sequence 30, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002

```

; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-30

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-199-637A-30 ..

Align seg 1/1 to: US-09-199-637A-30 from: 1 to: 798

1 GlyThr*****ProSer 6
|||||
3 GGCACATCATCAACG 20

seq_name: /cgn2_5/ptodata/2/ina/5A_COMB.seq:US-08-051-935A-49

seq_documentation_block:
; Sequence 49, Application US/08051935A
; Patent No. 5610050

; GENERAL INFORMATION:
; APPLICANT: Hubert E. Blum
; APPLICANT: Tsanyang Liang
; APPLICANT: Eithan Galun
; APPLICANT: Jack R. Wands
; TITLE OF INVENTION: Methods of Preventing
; TITLE OF INVENTION: Viral Replication
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/051,935A
; FILING DATE: 04/23/93
; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/846,328
; FILING DATE: 03/05/92
; APPLICATION NUMBER: 07/511,428
; FILING DATE: 04/20/90
; ATTORNEY/AGENT INFORMATION:
; NAME: Janis K. Fraser
; REGISTRATION NUMBER: 34,819

; REFERENCE/DOCKET NUMBER: 00786/185001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-051-935A-49

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-08-051-935A-49 ..

Align seg 1/1 to: US-08-051-935A-49 from: 1 to: 816

1 GlyThr*****ProSer 6
|||||
628 GGGACTTCACCCCGTCA 645

seq_name: /cgn2_5/ptodata/2/ina/6B_COMB.seq:US-09-020-956-14

seq_documentation_block:
; Sequence 14, Application US/09020956
; Patent No. 6261562

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-020-956-14

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-020-956-14/rev ..

Align seg 1/1 to reverse of: US-09-020-956-14 from: 1 to: 816

1 GlyThr*****ProSer 6

```

|||||
85 GGTACTACACCCCTTCA 68

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-607-14

seq_documentation_block:
; Sequence 14, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION/DOCKET NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-030-607-14

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-030-607-14/rev ..

Align seg 1/1 to reverse of: US-09-030-607-14 from: 1 to: 816

1 GlyThr*****ProSer 6
|||||
85 GGTACTACACCCCTTCA 68

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-607-14

seq_documentation_block:
; Sequence 14, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael

```

```

; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(816)
; OTHER INFORMATION: n = A,T,C or G
US-09-439-313-14

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-439-313-14/rev ..

Align seg 1/1 to reverse of: US-09-439-313-14 from: 1 to: 816

1 GlyThr*****ProSer 6
|||||
85 GGTACTACACCCCTTCA 68

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-361-707-104

seq_documentation_block:
; Sequence 104, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; Li, Jisu
; Wands, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 00786/287003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 104:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 821 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-361-707-104

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-361-707-104 ..
Align seg 1/1 to: US-09-361-707-104 from: 1 to: 821
1 GlyThr*****ProSer 6
|||||
581 GGGACTTCACCCCGTCA 598

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-434-774-5

seq_documentation_block:
; Sequence 5, Application US/09434774A
; Patent No. 6235514
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ISOPENTENYL
; TITLE OF INVENTION: MONOPHOSPHATE KINASE, AND METHODS OF USE
; FILE REFERENCE: wsr14448
; CURRENT APPLICATION NUMBER: US/09/434,774A
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(852)
US-09-434-774-5

alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x US-09-434-774-5/rev ..
Align seg 1/1 to reverse of: US-09-434-774-5 from: 1 to: 852
1 GlyThr*****ProSer 6
|||||
428 GGACATCTGGCCCGAC 411

```


OM of: BASK-853-CLAIM5 to: EST.* out_format : pfs

Date: Sep 24, 2002 12:52 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame.p2n.model -DEV=slp
-O/cn2.1/USPTO.spool/BASKAR853079/runat_17092002.103407_1808/app_query.fasta.1.98
-DB=EST -Qfmt=fastap -SUFFIX=mod.rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOPC=0.000 -LOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=50 -MODEALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=BASKAR853079 @CGNL.1.3956
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIME=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: BASK-853-CLAIM5

Query length: 6

Database: EST.*

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 2564.800000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_gss:TA62A100	+	22.00	112.75	4.9e+03	23
gb_est1:AA58120	+	22.00	106.96	1.0e+04	52
gb_est2:BI828531	+	22.00	106.96	1.0e+04	52
gb_est1:AA589681	+	22.00	106.96	1.1e+04	58
gb_gss:TA22040	+	22.00	105.19	1.3e+04	64
gb_gss:AA53886	+	22.00	105.38	1.3e+04	65
gb_est1:AA671552	+	22.00	104.08	1.5e+04	78
gb_est1:AU076593	+	22.00	103.82	1.6e+04	81
gb_est2:BE515109	+	22.00	103.07	1.7e+04	90
gb_gss:AZ433387	+	22.00	103.07	1.7e+04	90
gb_est1:AI022557	+	22.00	102.99	1.7e+04	91
gb_gss:AO822929	+	22.00	102.99	1.7e+04	91
gb_est1:AA864024	+	22.00	102.76	1.8e+04	94
gb_gss:CNS02A93	+	22.00	102.76	1.8e+04	94
gb_est1:AI142304	+	22.00	102.32	1.9e+04	99
gb_est2:BF833911	+	22.00	102.32	1.9e+04	100
gb_est1:AA721092	+	22.00	102.25	1.9e+04	101
gb_est2:BG342738	+	22.00	102.11	1.9e+04	103
gb_est2:BF481705	+	22.00	102.04	1.9e+04	104
gb_gss:BH342847	+	22.00	102.04	1.9e+04	104
gb_est1:AW066012	+	22.00	101.97	2.0e+04	105
gb_est1:AL508153	+	22.00	101.91	2.0e+04	106
gb_est2:BF082464	+	22.00	101.91	2.0e+04	106
gb_est1:AI920378	+	22.00	101.84	2.0e+04	107
gb_est2:BM377024	+	22.00	101.77	2.0e+04	108
gb_est1:AA996441	+	22.00	101.64	2.0e+04	110
gb_est2:HH39354	+	22.00	101.58	2.1e+04	111
gb_est1:AI040824	+	22.00	101.51	2.1e+04	112
gb_est2:BF871071	+	22.00	101.45	2.1e+04	113
gb_est2:BI047493	+	22.00	101.45	2.1e+04	113
gb_est2:BF872796	+	22.00	101.39	2.1e+04	114
gb_est1:AA003067	+	22.00	101.27	2.2e+04	116
gb_est1:AI964424	+	22.00	101.27	2.2e+04	116
gb_gss:BH196225	+	22.00	101.20	2.2e+04	117
gb_est2:BF202687	+	22.00	101.20	2.2e+04	117
gb_est2:BF287289	+	22.00	101.20	2.2e+04	117
gb_gss:BH050852	+	22.00	101.20	2.2e+04	117
gb_est1:AA773570	+	22.00	101.14	2.2e+04	118

gb_est2:BF199641	-	22.00	101.14	2.2e+04	118
gb_gss:AO983105	+	22.00	101.08	2.2e+04	119
gb_est2:BE700048	-	22.00	101.02	2.2e+04	120
gb_est1:AI020839	-	22.00	100.97	2.2e+04	121
gb_est1:AW847482	-	22.00	100.97	2.2e+04	121
gb_est1:AA065904	-	22.00	100.91	2.3e+04	122
gb_est2:BF856188	-	22.00	100.91	2.3e+04	122
gb_est2:BI777877	-	22.00	100.91	2.3e+04	122
gb_est1:AW878149	-	22.00	100.85	2.3e+04	123
gb_est2:N60943	+	22.00	100.85	2.3e+04	123

seq_name: gb_gss:TA62A100

seq_documentation_block:

LOCUS TA62A100 23 bp DNA linear GSS l3-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 62a10, reverse sequence,
genomic survey sequence.

ACCESSION AL464584.1 GI:11859285

VERSION AL464584

KEYWORDS GSS.

SOURCE Trypanosoma brucei.

ORGANISM Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

REFERENCE 1 (bases 1 to 23)

AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,

Melville, S.E., Rajandream, M.A. and Barrell, B.G.

TITLE Direct Submission

JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing

project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

nh@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TREGU927/4 GUTAT 10.1) was mechanically sheared

to give a tight size distribution (

4 kb). The + i method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small

insert libraries for whole genome shotgun sequencing projects. In

Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barrell, Oxford University Press, 1999).

Email: nelsaved@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available

at http://www.sanger.ac.uk/Projects/T-brucei/.

FEATURES

source 1..23

Location/Qualifiers

base COUNT 5 a 7 c 5 g 6 t

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6

Ratio: 3.667 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x TA62A100 ..

Align seg 1/1 to: TA62A100 from: 1 to: 23

1 GlyThr*****ProSer 6

|||||

3 GCACCATGTGCTCACT 20

seq_name: gb_est1:AA458120

seq_documentation_block:

LOCUS AA458120 52 bp mRNA linear EST 06-JUN-1997

seq_name: gb_est2:BI828531

REFERENCE JOURNAL
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 58)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

THE WASHU-HMMI MOUSE EST PROJECT
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:544248

Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
High quality sequence stop: 1.
Location/Qualifiers
source
1. .58
/organism="Mus musculus"
/strain="B6D2 Fl/J"
/db_xref="taxon:10090"
/clone="IMAGE:947392"
/clone_lib="Knowles Solter mouse blastocyst B1"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH1b"
/note="Organ: Embryo; Vector: pSPORT; Site_1: NotI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: SalI(dT): 5'-CGTGCACCGTCGACGTTTTTTTTTTT-3'. cDNAs were cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3."

BASE COUNT 14 a 20 c 15 g 9 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AA589681 ..
Align seg 1/1 to: AA589681 from: 1 to: 58

1 GlyThr*****ProSer 6
|||||
39 GGGACGGCAAGCCCATCC 56

seq_name: gb_gss:TA2F04Q

seq_documentation_block:
LOCUS TA2F04Q 64 bp DNA linear GSS 16-SEP-2000
DEFINITION T. brucei sheared genomic DNA clone 2f04, reverse sequence, genomic survey sequence.
ACCESSION AL441818
VERSION AL441818.1 GI:10185364
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
1 (bases 1 to 64)

REFERENCE 1
AUTHORS Hall, N., Bowman, S., Lennard, N. J., Doggett, J., Atkin, R.,

TITLE
JOURNAL

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S. E., Rajandream, M. A. and Barrell, B. G.
Direct Submission
Submitted (15-SEP-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk

COMMENT
Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v+1 method used for the library construction is described in detail in Smith, H. and Venter, J. C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES
Location/Qualifiers
source
1. .64
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="2f04"

BASE COUNT 11 a 18 c 20 g 15 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x TA2F04Q/rev ..
Align seg 1/1 to reverse of: TA2F04Q from: 1 to: 64

1 GlyThr*****ProSer 6
|||||
45 GGACCACTTCACCTCG 28

seq_name: gb_gss:AZ583866

seq_documentation_block:
LOCUS AZ583866 65 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0388103F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0388103 F, DNA sequence.

ACCESSION AZ583866
VERSION AZ583866.1 GI:11704178
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 65)

REFERENCE 1
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weis, R.

TITLE
JOURNAL

COMMENT
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```

Plate: 0388 row: I column: 03
Seq primer: CGTTGTAACAGCGGCAGT
Class: plasmid ends
High quality sequence stop: 65.
Location/Qualifiers
1. .65
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0388103"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/notes="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      16 a  17 c  12 g  20 t
ORIGIN
alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AZ583866 ..
Align seg 1/1 to: AZ583866 from: 1 to: 65

1 GlyThr*****ProSer 6
|||||
47 GGGACGAGTGCCCGCTCC 64

seq_name: gb_estl:AA671552

seq_documentation_block:
LOCUS      AA671552              78 bp  mRNA  linear  EST 25-NOV-1997
DEFINITION vn94h02.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
IMAGE:1039635 5', mRNA sequence.
ACCESSION  AA671552
VERSION    AA671552.1 GI:2643631
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 78)
REFERENCE  Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
AUTHORS   Geisel, S., Kuback, T., Lacy, M., Le, M., Martin, J., Morris, M.,
            Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
            Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
            Waterston, R.
            The WashU-HMI Mouse EST Project
            Unpublished (1996)
            Contact: Marra M/Mouse EST Project
            WashU-HMI Mouse EST Project

Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMGE Consortium (info@image.llnl.gov) for further information.
MGI:579659
Seq primer: -28ml3 rev1 ET from Amersham.
FEATURES
    source
    1..78
    Location/Qualifiers
    /organism="Mus musculus"
    /strain="NIH/Swiss"
    /db_xref="taxon:10090"
    /clone="IMAGE:1039635"
    /clone_lib="Stratagene mouse heart (#937316)"
    /sex="pooled"
    /tissue_type="heart"
    /dev_stage="13 day embryos"
    /lab_host="SOLR (kanamycin resistant)"
    /notes="Organ: heart; Vector: pBluescript SK-; Site:1:
    EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
    Oligo dT, 93 pooled NIH/Swiss 13 day embryo hearts.
    Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
    adaptor sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor
    sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT      16 a  23 c  32 g  7 t
ORIGIN
alignment_scores:
  Quality:      22.00      Length:      6
  Ratio:        3.667      Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AA671552/rev ..
Align seg 1/1 to reverse of: AA671552 from: 1 to: 78

1 GlyThr*****ProSer 6
|||||
68 GGAACGAGCTCCCTTCA 51

seq_name: gb_estl:AU076593

seq_documentation_block:
LOCUS      AU076593              81 bp  mRNA  linear  EST 04-MAY-2000
DEFINITION AU076593 Sugano cDNA library Homo sapiens cDNA clone FATB0009
similar to 5'-end region of H.sapiens dermatopontin mRNA, mRNA
sequence.
ACCESSION  AU076593.1 GI:7439027
VERSION    AU076593
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 81)
REFERENCE  Suzuki, Y., Ishihara, D., Sasaki, M., Nakagawa, H., Hata, H., Tsunoda, T.,
AUTHORS   Suzuki, Y., Ishihara, D., Sasaki, M., Nakagawa, H., Hata, H., Tsunoda, T.,
            Watanabe, M., Komatsu, T., Ota, T., Isogai, T., Suyama, A. and Sugano
            , S.
            Statistical analysis of the 5' untranslated region of human mRNA
            using 'Oligo-Capped' cDNA libraries
            Genomics 64 (3), 286-297 (2000)
            20221373
            Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
            , S. Construction and characterization of a full length-enriched and

```

a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
This clone was obtained from a 'full length-enriched' cDNA library
constructed by 'Oligo-Capping' method. The coding region starts
from the 50 bp upstream to the 3'-end.

FEATURES

source
1. .81
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FATB0009"
/clone_lib="Sugano cDNA library"
14 a 25 c 20 g 22 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AU076593 ..

Align seg 1/1 to: AU076593 from: 1 to: 81

1 GlyThr*****Proser 6

|||||
52 GGTACTTCTGCCCTAGT 69

seq_name: gb_est2:BE515109

seq_documentation_block:

LOCUS BE515109 90 bp mRNA linear EST 07-AUG-2000
DEFINITION 601236196F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3608326 5',
mRNA sequence.

ACCESSION BE515109
VERSION BE515109.1 GI:9722324
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM260 row: p column: 23.

Location/Qualifiers

1. .90

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3608326"

/clone_lib="NIH_MGC_44"

/tissue_type="endometrium, adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pORF7; Site_1: XhoI; Site_2:

ECORI; cDNA made by oligo-dT priming. Directionally

cloned into EcORI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 20 a 17 c 44 g 9 t

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BE515109/rev ..

Align seg 1/1 to reverse of: BE515109 from: 1 to: 90

1 GlyThr*****Proser 6

|||||
53 GGTACCACTGCGCCCTCC 36

seq_name: gb_gss:AZ433387

seq_documentation_block:

LOCUS AZ433387 90 bp DNA linear GSS 03-OCT-2000
DEFINITION 1M0219A15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0219A15 F, DNA sequence.

ACCESSION AZ433387

VERSION AZ433387.1 GI:10557400

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0219 row: A column: 15

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 90.

Location/Qualifiers

1. .90

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0219A15"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: pWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 24 a 34 c 12 g 20 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AZ433387 ..

Align seg 1/1 to: AZ433387 from: 1 to: 90

1 GlyThr*****ProSer 6
|||||
1 GGCACCAACCCCTTCC 18

seq_name: gb_est1:AI022557

seq_documentation_block:

LOCUS AI022557 91 bp mRNA linear EST 28-AUG-1998
DEFINITION O58e10.s1 Soares_parathyroid_tumor_NDHFA Homo sapiens cDNA clone IMAGE:1651050 3' similar to gb:X12791 19 KD PROTEIN OF SIGNAL RECOGNITION PARTICLE (HUMAN); mRNA sequence.

ACCESSION AI022557
VERSION AI022557.1 GI:3237798
KEYWORDS EST.
SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 91)

REFERENCE

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo

, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 883 Std Error: 0.00

Seq primer: -40m13 fwd. Et from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1..91

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1651050"

/clone_lib="Soares_parathyroid_tumor_NbHFA"

/tissue_type="parathyroid tumor"

/dev_stage="adult"

/lab_host="DH10B (ampicillin resistant)"

/note="organ: parathyroid gland; Vector: pTT73D (Pharmacia

) with a modified polylinker; Site_1: Not I; Site_2: Eco

Ri; lstr strand cDNA was primed with a Not I - oligo(dT)

primer

[5'-TGTTACCAATCTGAAGTGGGCGCGCCACCAATTTTTTTTTTTTTTTT

TTTTT-3'], double-stranded cDNA was size selected, ligated

to Eco RI adapters (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of a modified pTT73

vector (Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid

adenomas was kindly provided by Dr. Stephen Marx, National Institute of Diabetes and Digestive and Kidney Diseases, NIH."

BASE COUNT 25 a 19 c 23 g 23 t 1 others
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AI022557/rev ..

Align seg 1/1 to reverse of: AI022557 from: 1 to: 91

1 GlyThr*****ProSer 6
|||||
82 GGGACCGCGCTCCGCTCT 65

seq_name: gb_gss:AQ852929

seq_documentation_block:

LOCUS AQ852929 91 bp DNA linear GSS 25-MAY-2001
DEFINITION LMAJFV1_lm93g10.y1 Leishmania major FV1 random genomic library Leishmania major genomic clone LMAJFV1_lm93g10 5', similar to TR:060975 060975 L3162-2. ; contains element V1-ch1_type_II.5 leishmania repetitive element ; DNA sequence.

ACCESSION AQ852929
VERSION AQ852929.1 GI:6119254
KEYWORDS GSS.

SOURCE Leishmania major.

ORGANISM

Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;

Leishmania.

1 (bases 1 to 91)

Akopyants,N.S., Clifton,S.W., Martin,J., Pape,D., Wylie,T., Li,L.,

Kissinger,J., Roos,D.S., Marra,M., Hillier,L., Chinwalla,A.,

Blistain,A., Schmitt,A., Person,B., Theising,B., Ritter,E., Ronko

,I., Bennett,J., Cole,R., Underwood,K., Cardenas,M., Gibbons,M.,

Harvey,N., McCann,R., Tsagareishvili,R., Williams,T., Jackson,Y.,

Bowers,Y., Swaller,T., Waterston,R., Wilson,K. and Beverley,S.M.

A survey of the Leishmania major Friedlin strain V1 genome by

shotgun sequencing: a resource for DNA microarrays and expression

profiling

Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)

21192569

Contact: Akopyants, NS / Beverley, SM

Washu Leishmania Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewartson.wustl.edu

Library construction: Natalia S. Akopyants, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

If using this information please cite:

N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major

Friedlin strain V1 genome by shotgun sequencing' and the Washington

University Genome Sequencing Center For information on obtaining

clone material please contact: Natalia S. Akopyants Ph.D.

(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.

(beverley@borcim.wustl.edu)

Seq primer: -40RP from Gibco

Class: shotgun

High quality sequence stop: 82.

Location/Qualifiers

1..91

/organism="Leishmania major"

/strain="Friedlin strain V1"

/db_xref="taxon:5664"

/clone="LMAJFV1_lm93g10"

FEATURES
source

```
alignment_block:
  BASK-853-CLAIM5 x AA864024/rev ..
  Align seq 1/1 to reverse of: AA864024 from: 1 to: 94
```

Align seg 1/1 to reverse of: AA864024 from: 1 to: 94

1 GlyThr*****ProSer 6
|||||
33 GGCACGACGAGTCTCTAGC 16

33 GGCACCAGCAGTCCTTA
seq name: qb gss:CNS02A9J

seq_name: gb_gss:CNS02A9J	seq_documentation_block: LOCUS CNS02A9J Tetradion nigroviridis genome survey PUC-Ori end of clone DEFINITION 251802 of library G from Tetradion nigroviridis, genomic survey sequence.	94 bp DNA linear	GSS 12-MAY-2000
---------------------------	--	------------------------	-----------------

DEFINITION

Tetraodon nigroviridis
2S1E02 of library G from
sequence.

ACCESSION
VERSION ALI88272.1 GI:7826376
KEYWORDS GSS; genome survey seq
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis

accession number G817020370
 GSS: genome survey sequence.
 Tetraodon nigroviridis.
 Tetraodon nigroviridis
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 94)
 Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach, J.

AUTHORS	Roest-Crollius, H., Jaillon, O., Dasilva, C., Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 94)
AUTHORS	Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fischer, C., Bernot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Unpublished

REFERENCE	3 (bases 1 to 94)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .

COMMENT
submitted (12 APR 2007) to the EMBL/GenBank/EBI databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES	Source
1-94	Location/Qualifiers
	/organism="Tetraodon nigroviridis"
	/db_xref="taxon:99883"
	/clone="251E02"
	/clone_lib="G"
	/note="Genoscope sequence ID : COAG251BC01SP1-end ;

BASE COUNT 29 a 24 c 27 g 12 t 2 others
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x CNS02A9J ..

Align seg 1/1 to: CNS02A9J from: 1 to: 94

1 GlyThr*****ProSer 6
|||||
36 GGCACCTCCACACCTCT 53

seq_name: gb_est2:BG036840

seq_documentation_block:
LOCUS BG036840 99 bp mRNA linear EST 24-JAN-2001
DEFINITION 602326778F1 NIH_MGC_91 Homo sapiens CDNA clone IMAGE:4428426 5',
mRNA sequence.

ACCESSION BG036840.1 GI:12432267

VERSION BG036840
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 99)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DTp

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10178 row: k column: 19

High quality sequence stop: 99.

FEATURES

source

1..99
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4428426"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 20 a 28 c 38 g 13 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BG036840/rev ..

Align seg 1/1 to reverse of: BG036840 from: 1 to: 99

1 GlyThr*****ProSer 6
|||||
59 GGCACGAGCTCACCTCC 42

seq_name: gb_est1:A1142304

seq_documentation_block:

LOCUS A1142304 100 bp mRNA linear EST 25-SEP-1998
DEFINITION gg62f09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1839785
5', similar to gb:L19686_rnal MACROPHAGE MIGRATION INHIBITORY FACTOR
(HUMAN); mRNA sequence.

ACCESSION A1142304

VERSION A1142304.1 GI:3658663

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 100)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo,
Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -28ml3 rev2 ET from AmerSham

High quality sequence stop: 1.

FEATURES

source

1..100
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1839785"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTATACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 17 a 21 c 31 g 31 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x A1142304/rev ..

Align seg 1/1 to reverse of: A1142304 from: 1 to: 100

1 GlyThr*****ProSer 6

|||||

33 GGAAACAGTACACCTTCA 16

LOCUS	103 bp	mRNA	linear	EST 25-MAY-2001
DEFINITION	352627	MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.		
LOCUS	BG834414			

KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 103)
Fahrenkrug,S.C., Preking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keeler,J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine

JOURNAL Unpublished (2000)

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCGCCAGTCACGACG
Plate: 113 row: G column: 22
Seq primer: ATTAGTGACACTAG.

FEATURES
source
1..103
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC IPIC"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 23 a 26 c 27 g 27 t

ORIGIN
|||||
1 GlyThr*****ProSer 6
40 GGAACCTTCGACGCCATCA 23

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BG34414/rev ..

Align seg 1/1 to reverse of: BG34414 from: 1 to: 103

seq_name: gb_est2:BG342738

seq_documentation_block:
LOCUS BG342738 104 bp mRNA linear EST 27-FEB-2001
DEFINITION 602374558F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4481991 5', mRNA sequence.

ACCESSION BG342738

VERSION BG342738.1 GI:13149176

KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 104)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10318 row: c column: 16
High quality sequence stop: 104.

FEATURES
source
1..104
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGE:4481991"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 41 a 28 c 18 g 17 t

ORIGIN
|||||
1 GlyThr*****ProSer 6
34 GGCACAACTCTCGCTCC 51

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BG342738 ..

Align seg 1/1 to: BG342738 from: 1 to: 104

seq_name: gb_est2:BF481705

seq_documentation_block:
LOCUS BF481705 104 bp mRNA linear EST 05-DEC-2000
DEFINITION FM1_21_D04_b1_A003 Floral-Induced Meristem 1 (FM1) Sorghum propinquum cDNA, mRNA sequence.

ACCESSION BF481705

VERSION BF481705.1 GI:11552526

KEYWORDS EST.

SOURCE Sorghum propinquum.
ORGANISM Sorghum propinquum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
AUTHORS Cordonnier-pratt,M.-M., Gingle,A., Sudman,M., Marsala,C. and Pratt,L.H.
TITLE An EST database from Sorghum: floral-induced meristems
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence start: 76
High quality sequence stop: 100
POLYA-No.

FEATURES
source

Location/Qualifiers
1. .104
/organism="Sorghum propinquum"
/db_xref="taxon:132711"
/clone_lib="Floral-Induced Meristem 1 (FMI)"
/note="Organ: Floral-Induced meristems; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."
19 a 39 c 29 g 17 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BF481705 ..

Align seg 1/1 to: BF481705 from: 1 to: 104

1 GlyThr*****ProSer 6

|||||

38 GGTACTGCCTCACCCTCC 55

seq_name: gb_gss:BH342847

seq_documentation_block:

LOCUS BH342847 104 bp DNA linear GSS 03-DEC-2001
DEFINITION CH230-41L13.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-41L13, DNA sequence.

ACCESSION BH342847

VERSION BH342847.1 GI:17273581

KEYWORDS GSS.

SOURCE Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 104)

REFERENCE
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
.A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)

Other_GSSs: CH230-41L13.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel.: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or_ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 41 row: L column: 13

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. .104
/organism="Rattus norvegicus"

/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-41L13"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

BASE COUNT 31 a 26 c 20 g 27 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BH342847 ..

Align seg 1/1 to: BH342847 from: 1 to: 104

1 GlyThr*****ProSer 6

|||||

79 GGCACACACACCAAGC 96

seq_name: gb_est1:AW066012

seq_documentation_block:

LOCUS AW066012 105 bp mRNA linear EST 30-MAR-2000
DEFINITION 687004F12.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
sequence.

ACCESSION AW066012.1 GI:6021084

VERSION AW066012

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 105)

AUTHORS Walbot,V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 687004 row: F column: 12.

FEATURES

source

Location/Qualifiers
1. .105
/organism="Zea mays"
/cultivar="Illinois High Oil"
/db_xref="taxon:4577"
/clone_lib="687 - Early embryo from Delaware"
/tissue_type="embryo"
/dev_stage="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"

/note="Organ: embryo; Vector: pBluescript SK; Site_1: XhoI
; Site_2: EcoRI; Library was prepared by Statagene using
the Uni-ZAP XR system (Stratagene BN937328-12). Clones
were picked by a Q-bot after blue/white selection
(ampicillin resistance - use 100 micrograms/microliter).

Developed from a pool of equal amounts of RNA from
developing embryos sampled at 14, 21, 28 and 35 days after
pollination of the Illinois High Oil Maize strain Cycle
90. This closed strain has been selected for high oil
concentration for 90 generations and originates from the

```
17 a 47 c 31 g 10 t
1890s era open pollinated variety Burr's White"
BASE COUNT
ORIGIN

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AW066012 ..
Align seg 1/1 to: AW066012 from: 1 to: 105

1 GlyThr*****ProSer 6
|||||
31 GGCACGACGTCGCGAGT 48

seq_name: gb_est1:AL508153

seq_documentation_block:
LOCUS AL508153 106 bp mRNA linear EST 04-JAN-2001
DEFINITION Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
ACCESSION AL508153
VERSION AL508153.1 GI:12034368
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 106)
Michalek,W., Weschke,W., Pleissner,K.-P. and Graner,A.
EST sequencing and analysis in barley
Unpublished (2000)
Contact: Michalek W
Institute for Plant Genetics and Crop Plant Research
Corrensstr.3, D-06466 Gatersleben, Germany
Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
Seq primer: T3 primer for 5'end.
Location/Qualifiers
1..106
/organism="Hordeum vulgare"
/cultivar="Barke"
/db_xref="taxon:4513"
/clone="HY07024V"
/clone_lib="Hordeum vulgare Barke developing caryopsis
(3.-15.DAP)"
/tissue_type="developing caryopsis (3.-15.DAP)"
/lab_host="XLOR"
/note="Vector: plasmid pBK-CMV; Site_1: EcoRI; Site_2:
XhoI; mRNA was made from developing caryopsis (3.-15.DAP)
of spring barley variety 'Barke', a high quality malting
variety. Cloning sites: EcoRI (5'-end of cDNA) and XhoI
(3'-end of cDNA). NOTE: Due to a cloning artefact caused
by the kit, in most cases the EcoRI site is NOT present,
as well as the EcoRI adapter. Average insert size is 1 kb
Sequence trimming: Vector sequences and sequence ends were
trimmed from the 5'- and 3'-end until a 50 bp window
contains less than two ambiguities. The maximum length was
set to 700 bp"
11 a 44 c 34 g 17 t
BASE COUNT
ORIGIN

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AW066012 ..
Align seg 1/1 to reverse of: BASK-853-CLAIM5 from: 1 to: 106
```

```
1 GlyThr*****ProSer 6
|||||
42 GGCACGACGTCGCGTCG 25

seq_name: gb_est2:BF082464

seq_documentation_block:
LOCUS BF082464 106 bp mRNA linear EST 18-OCT-2000
DEFINITION RC3-BF0333-110900-116-dll_1 Homo sapiens cDNA, mRNA
sequence.
ACCESSION BF082464
VERSION BF082464.1 GI:10876294
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 106)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&l2=RC3-BF0333-110
900-116-dll_1&t3=2000-09-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 106.
Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BF0333"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
21 a 26 c 31 g 28 t
BASE COUNT
ORIGIN

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BF082464/rev ..
Align seg 1/1 to reverse of: BF082464 from: 1 to: 106
```

```

1 GlyThr*****ProSer 6
|||||
57 GCACAGTACACCTTCA 40

seq_name: gb_est1:AI920378

seq_documentation_block:
LOCUS AI920378 107 bp mRNA linear EST 29-JUL-1999
DEFINITION 603020D10.x1 603 - stressed root cDNA library from Wang/Bohnert lab
ACCESSION AI920378
VERSION AI920378.1 GI:5650018
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
JOURNAL clade: Panicoideae; Andropogoneae; Zea.
COMMENT 1 (bases 1 to 107)
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603020 row: D column: 10.
FEATURES
source
1..107
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from
Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+) XR;
Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 23 a 27 c 31 g 26 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AI920378 ..
Align seg 1/1 to: AI920378 from: 1 to: 107

1 GlyThr*****ProSer 6
|||||
20 GGAAGTGCATCTCCTCG 37

seq_name: gb_est2:BM377024

seq_documentation_block:
LOCUS BM377024 108 bp mRNA linear EST 10-JAN-2002
DEFINITION Ebem05_SQ003_O10_R IGF Barley Ebem05 library Hordeum vulgare cDNA
ACCESSION BM377024
VERSION BM377024.1 GI:18120414
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Zea mays.
JOURNAL clade: Panicoideae; Andropogoneae; Zea.
COMMENT 1 (bases 1 to 107)
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603020 row: D column: 10.
FEATURES
source
1..107
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from
Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+) XR;
Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 23 a 27 c 31 g 26 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AI920378 ..
Align seg 1/1 to: AI920378 from: 1 to: 107

1 GlyThr*****ProSer 6
|||||
20 GGAAGTGCATCTCCTCG 37

seq_name: gb_est2:BM377024

seq_documentation_block:
LOCUS BM377024 108 bp mRNA linear EST 10-JAN-2002
DEFINITION Ebem05_SQ003_O10_R IGF Barley Ebem05 library Hordeum vulgare cDNA
ACCESSION BM377024
VERSION BM377024.1 GI:18120414
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Zea mays.
JOURNAL clade: Panicoideae; Andropogoneae; Zea.
COMMENT 1 (bases 1 to 107)
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603020 row: D column: 10.
FEATURES
source
1..107
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from
Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+) XR;
Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 23 a 27 c 31 g 26 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BM377024/rev...
Align seg 1/1 to reverse of: BM377024 from: 1 to: 108

1 GlyThr*****ProSer 6
|||||
54 GGGACCGCAGCGCCGTCG 37

seq_name: gb_est1:AA896441

seq_documentation_block:
LOCUS AA896441 110 bp mRNA linear EST 06-APR-1998
DEFINITION vy3ig10.r1 Stratagene mouse lung 937302 Mus musculus cDNA clone
IMAGE:1297122 5' similar to gb:M21532 Mus musculus PCD-5 mRNA,
complete cds (MOUSE);, mRNA sequence.
ACCESSION AA896441
VERSION AA896441.1 GI:3032834
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 110)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

```

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 108)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: twaugh@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.
FEATURES
source
1..108
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone_lib="EBem05_SQ003_O10"
/tissue_type="Embryo"
/dev_stage="14 days post anthesis"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from developing grains (14
days post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT 12 a 41 c 39 g 16 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BM377024/rev...
Align seg 1/1 to reverse of: BM377024 from: 1 to: 108

1 GlyThr*****ProSer 6
|||||
54 GGGACCGCAGCGCCGTCG 37

seq_name: gb_est1:AA896441

seq_documentation_block:
LOCUS AA896441 110 bp mRNA linear EST 06-APR-1998
DEFINITION vy3ig10.r1 Stratagene mouse lung 937302 Mus musculus cDNA clone
IMAGE:1297122 5' similar to gb:M21532 Mus musculus PCD-5 mRNA,
complete cds (MOUSE);, mRNA sequence.
ACCESSION AA896441
VERSION AA896441.1 GI:3032834
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 110)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

```

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:678170

Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 1.

FEATURES

source

1. .1110
 /organism="Mus musculus"
 /strain="C57BL/6 x CBA"
 /db_xref="taxon:10090"
 /clone="IMAGE:1297122"
 /clone_lib="Stratagene mouse lung 937302"
 /sex="female"
 /tissue_type="lung"
 /dev_stage="6-8 month old"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: lung; Vector: pBluescript SK-; Site.1: ECORI
 ; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo
 dt. 6-8 month old female lung and 1.5 year old male lung
 were source of mRNA. Average insert size: 1.5 kb; Uni-ZAP
 XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'
 adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT
 ORIGIN

20 a 23 c 41 g 26 t

alignment_scores:

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AA896441/rev ..

Align seg 1/1 to reverse of: AA896441 from: 1 to: 110

1 GlyThr*****ProSer 6

|||||

25 GGAACAGCAGCCCAAGC 8

seq_name: gb_est2:H39354

seq_documentation_block:

LOCUS H39354 111 bp mRNA linear EST 27-AUG-1996
 DEFINITION DR38 IFNgamma-induced astrocyte ESTs Rattus norvegicus cDNA clone
 DR38 5'-end with RP similar to unknown sequence, mRNA sequence.

ACCESSION H39354.1 GI:915028

VERSION H39354.1

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 111)

AUTHORS Kuchinke,W., Hart,R.P. and Jonakait,G.M.

TITLE Identification of mRNAs regulated by interferon-gamma in cultured
 rat astrocytes by PCR differential display

JOURNAL Neuroimmunomodulation 2, 347-355 (1995)

MEDLINE 96437728

COMMENT Contact: Hart RP

Rutgers University

Dept of Biol Sci, Rutgers University, 101 Warren St, Newark NJ

07102

Tel: 2016481008

Fax: 2016481007

Email: rhart@andromeda.rutgers.edu

unknown sequence

subcloned into PCRscript

Seq primer: RP.

FEATURES

source

1. .111
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="DR38"
 /clone_lib="IFNgamma-induced astrocyte ESTs"
 /note="CDNA from IFNgamma-induced cultured rat astrocytes
 was screened by PCR differential display. Regulated cDNAs
 were amplified, sequenced and confirmed by dot-blot and
 Northern." Northern.

BASE COUNT 34 a 25 c 28 g 24 t
 ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x H39354 ..

Align seg 1/1 to: H39354 from: 1 to: 111

1 GlyThr*****ProSer 6

|||||

67 GGAATCTCCACCCCACT 84

seq_name: gb_est1:AI040824

seq_documentation_block:

LOCUS AI040824 112 bp mRNA linear EST 28-AUG-1998
 DEFINITION ox27d11.sl Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
 IMAGE:1657557 3' similar to TR:P97582 P97582 ANKYRIN ;, mRNA
 sequence.

ACCESSION AI040824.1 GI:3280018

VERSION AI040824

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 112)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality

Insert Length: 745 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 1.

FEATURES

source

1. .112
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1657557"
 /clone_lib="Soares_total_fetus_Nb2HF8_9w"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dt) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 23 a 36 c 34 g 19 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AI040824/rev ..

Align seg 1/1 to reverse of: AI040824 from: 1 to: 112

1 GlyThr*****ProSer 6
24 GGGACATCTGCTCCCTCC 7

seq_name: gb_est2:BF871071

seq_documentation_block:
LOCUS BF871071 113 bp mRNA EST 17-JAN-2001
DEFINITION CM0-ET0121-311000-658-a10 ET0121 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF871071
VERSION BF871071.1 GI:12261201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 113)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM0&t2=CM0-ET0121-311000-658-a10&t3=2000-10-31&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 113.

Location/Qualifiers
1. .113

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0121"
/dev_stage="Adult"
/note="Organ: lung tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

BASE COUNT 29 a 26 c 31 g 27 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BF871071 ..

Align seg 1/1 to: BF871071 from: 1 to: 113

1 GlyThr*****ProSer 6
74 GGCACCACTACTCCCTCT 91

seq_name: gb_est2:BI047493

seq_documentation_block:
LOCUS BI047493 113 bp mRNA linear EST 14-JUN-2001
DEFINITION MR4-ST0240-080101-024-a09 ST0240 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI047493
VERSION BI047493.1 GI:14454115
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 113)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR4&t2=MR4-ST0240-080101-024-a09&t3=2001-01-08&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 113.

Location/Qualifiers
1. .113

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0240"
/dev_stage="Adult"
/note="Organ: stomach; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

BASE COUNT 18 a 37 c 30 g 28 t
ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BI047493 ..

Align seg 1/1 to: BI047493 from: 1 to: 113

1 GlyThr*****ProSer 6

|||||
80 GGTACAGCGCTCCTCC 97

seq_name: gb_est2:BF872796

seq_documentation_block:

LOCUS BF872796 114 bp mRNA linear EST 17-JAN-2001
DEFINITION QV0-ET0149-011100-464-a04_1 ET0149 Homo sapiens cDNA, mRNA

sequence.

ACCESSION BF872796

VERSION BF872796.1 GI:12262926

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,
Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H.,
Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,
M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-ET0149-
011100-464-a04_1&t3=2000-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 114.

FEATURES

source

1. .114
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0149"
/dev_stage="Adult"

/note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 31 a 27 c 28 g 27 t 1 others

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BF872796/rev ..

Align seg 1/1 to reverse of: BF872796 from: 1 to: 114

1 GlyThr*****ProSer 6

|||||
102 GGAAGTACTACGCCGAGC 85

seq_name: gb_est1:AA003067

seq_documentation_block:

LOCUS AA003067 116 bp mRNA linear EST 19-JUL-1996
DEFINITION mg48d08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA

clone IMAGE:427023 5', mRNA sequence.

ACCESSION AA003067

VERSION AA003067.1 GI:1446662

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 116)

REFERENCE

AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:261575
Seq primer: ETPrimer
High quality sequence stop: 114.

FEATURES

source

1. .116
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:427023"

/clone_lib="Soares mouse embryo NbME13.5 14.5"

/sex="unknown"

/tissue_type="embryo"

/dev_stage="13.5-14.5dpc total fetus"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGGAAATTTTTTTTTTTTTTTTTT

T 3', on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 22 a 25 c 29 g 23 t 17 others

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0


```

Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AA003067
..
Align seg 1/1 to: AA003067 from: 1 to: 116

1 GlyThr*****ProSer 6
|||||
57 GGGACCGCGCCCTCT 74

seq_name: gb_est1:AI964424

seq_documentation_block:
LOCUS AI964424
DEFINITION 496010A10.x1 496 - stressed shoot cDNA library from Wang/Bohnert
lab Zea mays cDNA, mRNA sequence.
ACCESSION AI964424
VERSION AI964424
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
1 (bases 1 to 116)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496010 row: A column: 10.

FEATURES
source
1..116
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from
Wang/Bohnert lab"
/tissue_type="Seedling"
/dev_stage="salt stress"
/lab_host="E.Coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+) XR;
Wang/Bohnert"

BASE COUNT 19 a 42 c 34 g 21 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x AI964424
..
Align seg 1/1 to: AI964424 from: 1 to: 116

1 GlyThr*****ProSer 6
|||||
16 GGAACAGCTTCTCCATCT 33

seq_name: gb_gss:BH196225

seq_documentation_block:
LOCUS BH196225
DEFINITION TC3-63B11.TV TC3 Trypanosoma cruzi genomic clone TC3-63B11, DNA

116 bp DNA linear GSS 24-OCT-2001
Trypanosoma cruzi genomic clone TC3-63B11, DNA

Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BH196225/rev
..
Align seg 1/1 to reverse of: BH196225 from: 1 to: 116

1 GlyThr*****ProSer 6
|||||
83 GGCACAGCGCAGCCCTCG 66

seq_name: gb_est2:BF202687

seq_documentation_block:
LOCUS BF202687
DEFINITION WHE1777_A07_B132S Wheat pre-anthesis spike cDNA library Triticum
aestivum cDNA clone WHE1777_A07_B13, mRNA sequence.
ACCESSION BF202687
VERSION BF202687
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum

117 bp mRNA linear EST 06-NOV-2000
Wheat pre-anthesis spike cDNA library Triticum
aestivum cDNA clone WHE1777_A07_B13, mRNA sequence.

sequence.
BH196225
VERSION BH196225.1 GI:16364269
KEYWORDS GSS.
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi.
Trypanosoma cruzi.
Trypanosoma cruzi.
Trypanosoma; Kinetoplastida; Trypanosomatidae;
Eukaryota; Euglenozoa; Schizotrypanum.
REFERENCE
1 (bases 1 to 116)
Kluge,S., Edwards,K.E., Nilsson,D., Bontempi,E.J., Myler,P., Stuart
,K., Ghedin,E., El-Sayed,N.M. and Andersson,B.
Clustering and analysis of BAC-end and GSS sequences from
Trypanosoma cruzi
Unpublished (2001)
Other_GSSs: TC3-63B11.TP
Contact: Bjorn Andersson
Department of Genetics and Pathology
Uppsala University
Rudbeck Laboratory, SE-751 85, Uppsala, Sweden
Tel: 46 18 471 4107
Fax: 46 18 471 4808
Email: bjorn.andersson@genpat.uu.se
Clones are derived from the Trypanosoma cruzi CL-Brener BAC library
TC3. For clone availability, please contact Dr. Bjorn Andersson at
Uppsala University (bjorn.andersson@genpat.uu.se).
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..116
/organism="Trypanosoma cruzi"
/strain="CL Brener"
/db_xref="taxon:5693"
/clone_lib="TC3-63B11"
/note="Vector: pBelOAC11; Site_1: Hin dIII; Constructed
for Uppsala University by Marie-Christine Le Paslier in
the laboratory of Denis Le Paslier at the Centre d'Etude
du Polymorphisme Humain (CEPH), Paris, France. Briefly,
Trypanosoma cruzi CL-Brener agarose embedded DNA (obtained
from Dr. Franco da Silveira) was partially digested with
Hin dIII. High molecular weight fragments were ligated in
pBelOAC11 digested with Hin dIII. The average insert
size is 100 kb. Total clone coverage: approx. 33 X the
haploid genome."
BASE COUNT 21 a 23 c 47 g 25 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
BASK-853-CLAIM5 x BH196225/rev
..
Align seg 1/1 to reverse of: BH196225 from: 1 to: 116

1 GlyThr*****ProSer 6
|||||
83 GGCACAGCGCAGCCCTCG 66

seq_name: gb_est2:BF202687

seq_documentation_block:
LOCUS BF202687
DEFINITION WHE1777_A07_B132S Wheat pre-anthesis spike cDNA library Triticum
aestivum cDNA clone WHE1777_A07_B13, mRNA sequence.
ACCESSION BF202687
VERSION BF202687
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum

117 bp mRNA linear EST 06-NOV-2000
Wheat pre-anthesis spike cDNA library Triticum
aestivum cDNA clone WHE1777_A07_B13, mRNA sequence.

```

```

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Pooideae
; Triticeae; Triticum.
1 (bases 1 to 117)
Anderson,O.D., Choi,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105959773
Fax: 5105959818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
    source
        1. .117
            /organism="Triticum aestivum"
            /cultivar="Chinese Spring"
            /db_xref="taxon:4565"
            /clone="WHE1777_A07_B13"
            /tissue_type="Spike before anthesis"
            /dev_stage="Adult plant"
            /lab_host="E. coli SOLR"
            /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
            Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
            greenhouse. Whole spike with awns trimmed, white, green
            and yellow anther were collected and total RNA, and
            poly(A) RNA were prepared, a cDNA library was made, and
            the cDNA clones were in vivo excised to give pBluescript
            phagemids in the TJ Close lab (Choi, Close, Fenton) at
            the University of California, Riverside. Plasmid DNA
            preparations and DNA sequencing were performed in the OD
            Anderson lab (all other authors)."
BASE COUNT      20 a   40 c   42 g   15 t
ORIGIN
alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x BF202687
..
Align seg 1/1 to: BF202687 from: 1 to: 117
1 GlyThr*****ProSer 6
|||||
13 GGGACATCGACTCCCAAGT 30
seq_name: gb_est2:BF287289
seq_documentation_block:
LOCUS      BF287289
DEFINITION EST451866 Rat Gene Index, normalized rat, Rattus norvegicus cDNA
Rattus norvegicus cDNA clone RGIGD88, mRNA sequence.
ACCESSION  BF287289
VERSION     BF287289.1 GI:11218345
KEYWORDS   EST.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE  1 (bases 1 to 117)
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 117)
Anderson,O.D., Choi,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105959773
Fax: 5105959818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
    source
        1. .117
            /organism="Triticum aestivum"
            /cultivar="Chinese Spring"
            /db_xref="taxon:4565"
            /clone="WHE1777_A07_B13"
            /tissue_type="Spike before anthesis"
            /dev_stage="Adult plant"
            /lab_host="E. coli SOLR"
            /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
            Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
            greenhouse. Whole spike with awns trimmed, white, green
            and yellow anther were collected and total RNA, and
            poly(A) RNA were prepared, a cDNA library was made, and
            the cDNA clones were in vivo excised to give pBluescript
            phagemids in the TJ Close lab (Choi, Close, Fenton) at
            the University of California, Riverside. Plasmid DNA
            preparations and DNA sequencing were performed in the OD
            Anderson lab (all other authors)."
BASE COUNT      20 a   40 c   42 g   15 t
ORIGIN
alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x BF202687
..
Align seg 1/1 to: BF202687 from: 1 to: 117
1 GlyThr*****ProSer 6
|||||
13 GGGACATCGACTCCCAAGT 30
seq_name: gb_est2:BF287289
seq_documentation_block:
LOCUS      BF287289
DEFINITION EST451866 Rat Gene Index, normalized rat, Rattus norvegicus cDNA
Rattus norvegicus cDNA clone RGIGD88, mRNA sequence.
ACCESSION  BF287289
VERSION     BF287289.1 GI:11218345
KEYWORDS   EST.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE  1 (bases 1 to 117)

```

AUTHORS

TITLE

JOURNAL

COMMENT

Malek,R.L., Cho,J., Lee,Y., Karamycheva,S., Parvizi,B., Pertea,G.,
Sultana,R., Tsai,J., White,J., Quackenbush,J. and Lee,N.H.
Generation of ESTs from Normalized Rat Embryo, Bento Soares
Unpublished (2000)
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org

This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information.

FEATURES

source

1. .117
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RGIGD88"
/clone_lib="Rat Gene Index, normalized rat, Rattus
norvegicus cDNA"
/tissue_type="mixed tissue"
/lab_host="DH5-alpha"
/note="Vector: pT3n7Pac; Site_1: EcoRI; Site_2: NotI;
Combination of ROV, RBR, RKI, RLI, RPL, RLU, REM, RMO, RSP
, RHE, RPC, RPN"
15 a 45 c 24 g 33 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BF287289 ..

Align seg 1/1 to: BF287289 from: 1 to: 117

1 GlyThr*****ProSer 6

|||||

80 GGGACTACCTCTCTCTCA 97

seq_name: gb_gss:BH050852

seq_documentation_block:

LOCUS BH050852
DEFINITION RPCI-24-23209.TV RPCI-24 Mus musculus genomic clone RPCI-24-23209,
DNA sequence.
ACCESSION BH050852
VERSION BH050852.1 GI:14841622
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 117)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 232 row: 0 column: 9
Seq primer: T7
Class: BAC ends.

FEATURES

Source
Location/Qualifiers
1. .117
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-23209"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 20 a 26 c 46 g 25 t
ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BH050852/rev ..

Align seg 1/1 to reverse of: BH050852 from: 1 to: 117

1 GlyThr*****ProSer 6

|||||

34 GCACATCTCTCCCTCT 17

seq_name: gb_est1:AA773570

seq_documentation_block:

LOCUS AA773570 118 bp mRNA linear EST 29-JAN-1998
DEFINITION ab61c01.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone IMAGE:845280 3', mRNA sequence.

ACCESSION AA773570.1 GI:2825141

VERSION AA773570

KEYWORDS EST.

SOURCE

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 118)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,

Krizman,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin

,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,

White,Y., Wyllie,T., Waterston,R. and Wilson,R.

WashU-NCI human EST Project

Unpublished (1997)

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 94.

Location/Qualifiers

1. .118

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:845280"

/clone_lib="Stratagene lung carcinoma 937218"

/tissue_type="lung carcinoma"

/cell_line="NCI-H69"

FEATURES

Source

1. .118

/organism="Triticum monococcum"

/cultivar="DV92"

/db_xref="taxon:4568"

/clone="WHE0591-0594_L23_L23"

/clone_lib="Triticum monococcum vegetative apex cDNA library"

/tissue_type="Vegetative shoot apex"

/dev_stage="Three weeks-old plants"

/lab_host="E. coli XL0LR"

/note="Vector: Lambda pBK-CMV (Lambda Zap Express), excised phagemid; Site 1: EcoRI; Site 2: XhoI; The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised at the

/dev_stage="cell line NCI-H69"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 17 a 43 c 35 g 23 t
ORIGIN

alignment_scores:

Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AA773570/rev ..

Align seg 1/1 to reverse of: AA773570 from: 1 to: 118

1 GlyThr*****ProSer 6

|||||

21 GGACCGCTACCCGCTCG 4

seq_name: gb_est2:BF199641

seq_documentation_block:

LOCUS BF199641 118 bp mRNA linear EST 16-APR-2001
DEFINITION WHE0591-0594_L23_L23ZE Triticum monococcum vegetative apex cDNA library Triticum monococcum cDNA clone WHE0591-0594_L23_L23, mRNA sequence.

ACCESSION BF199641

VERSION BF199641.1 GI:111113507

KEYWORDS EST.

SOURCE Triticum monococcum.

ORGANISM Triticum monococcum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 118)

AUTHORS Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia

C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,

Stamove,B. and Tong,J.C.

The structure and function of the expressed portion of the wheat

genomes - Vegetative apex cDNA library from Triticum monococcum

Unpublished (2001)

CONTACT: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: Stratagene T3 primer.

Location/Qualifiers

1. .118

/organism="Triticum monococcum"

/cultivar="DV92"

/db_xref="taxon:4568"

/clone="WHE0591-0594_L23_L23"

/clone_lib="Triticum monococcum vegetative apex cDNA library"

/tissue_type="Vegetative shoot apex"

/dev_stage="Three weeks-old plants"

/lab_host="E. coli XL0LR"

/note="Vector: Lambda pBK-CMV (Lambda Zap Express), excised phagemid; Site 1: EcoRI; Site 2: XhoI; The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised at the

University of California, Davis (V. Echenique, B. Stamova, J. Dubcovsky). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 24 a 33 c 35 g 26 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BF199641/rev ..

Align seg 1/1 to reverse of: BF199641 from: 1 to: 118

1 GlyThr*****ProSer 6
|||||
20 GGACCTCAGCGCCCTCG 3

seq_name: gb_gss:AQ983105

seq_documentation_block:

LOCUS AQ983105 119 bp DNA linear GSS 30-JAN-2000
DEFINITION RPCI-23-262P13.TJ RPCI-23 Mus musculus genomic clone RPCI-23-262P13
DNA sequence.

ACCESSION AQ983105
VERSION AQ983105.1 GI:6816310
KEYWORDS GSS.
SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 119)

Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html

Plate: 262 row: P column: 13

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..119

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-262P13"

/clone_lib="RPCI-23"

/sex="female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:

EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcoRI and EcoRI Methylase. Size

selected DNA was cloned into the pBACe3.6 vector at the

EcoRI sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 29 a 26 c 28 g 36 t
ORIGIN

alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AQ983105 ..

Align seg 1/1 to: AQ983105 from: 1 to: 119

1 GlyThr*****ProSer 6
|||||
23 GGGACCTCAGTCTCCCTCA 40

seq_name: gb_est2:BE700048

seq_documentation_block:

LOCUS BE700048 120 bp mRNA linear EST 12-SEP-2000
DEFINITION MR0-NN0087-040600-012-b10 NN0087 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE700048

VERSION BE700048.1 GI:10087790

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 120)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.A.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-MR0-NN0087-040

600-012-b10&t3=2000-06-04&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 25

High quality sequence stop: 120.

FEATURES

Location/Qualifiers

1..120

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NN0087"

/dev_stage="Adult"

/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

20 a 42 c 30 g 28 t

BASE COUNT

ORIGIN

```

alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x BE700048 ..
  Align seg 1/1 to: BE700048 from: 1 to: 120

1 GlyThr*****ProSer 6
|||||
89 GGAACCTCTGCTCCATCT 106

seq_name: gb_est1:AI020839

seq_documentation_block:
LOCUS      AI020839      121 bp      mRNA      linear      EST 16-JUN-1998
DEFINITION ub05d11.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
IMAGE:1366101.5' similar to TR:014899 O14899 ZINC FINGER PROTEIN ;,
mRNA sequence.
ACCESSION  AI020839
VERSION    AI020839.1 GI:32351751
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 121)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
            The WashU-HMMI Mouse EST Project
            Unpublished (1996)
            Contact: Marra M/Mouse EST Project
            WashU-HMMI Mouse EST Project
            Washington University School of Medicinep
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:899321
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 1.
            Location/Qualifiers
                1..121
                /organism="Mus musculus"
                /strain="C57Bl/6J"
                /db_xref="taxon:10090"
                /clone="IMAGE:1366101"
                /clone_lib="Soares_mammary_gland_NbMMG"
                /sex="male"
                /tissue_type="mammary gland"
                /dev_stage="4 weeks"
                /lab_host="DH10B"
            /note="Organ: mammary gland; Vector: p7T3D-Pac (Pharmacia
            ) with a modified polylinker; Site_1: Not I; Site_2: Eco
            RI; 1st strand cDNA was primed with a Not I - oligo(dT)
            primer [5'
            TGTACCATCTGAGTGGAGCGCGCGAATGCTTTTTTTTTTTTTTTTTTTTTTT
            T 3']; double-stranded cDNA was ligated to Eco RI
            adaptors (Pharmacia), digested with Not I and cloned into
            the Not I and Eco RI sites of the modified p7T3 vector.
            RNA provided by Dr. Minoru KO, Wayne State Univ. Library
            constructed and normalized by Bento Soares and M.Fatima
            Bonaldo."
            BASE COUNT      28 a      28 c      43 g      22 t

FEATURES
  source
    1..121
    /organism="Mus musculus"
    /db_xref="taxon:9606"
    /clone_lib="CT0209"
    /dev_stage="Adult"
    /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
    SmaI; A mini-library was made by cloning products derived
    from ORESTES PCR (U.S. Letters Patent application No. 196
    ,716 - Ludwig Institute for Cancer Research) profiles
    into the puc 18 vector. Reverse transcription of tissue
    mRNA and cDNA amplification were performed under low
    stringency conditions."
    BASE COUNT      36 a      41 c      18 g      26 t
    ORIGIN
    alignment_scores:

```

```

ORIGIN
alignment_scores:
  Quality: 22.00      Length: 6
  Ratio: 3.667        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  BASK-853-CLAIM5 x AI020839/rev ..
  Align seg 1/1 to reverse of: AI020839 from: 1 to: 121

1 GlyThr*****ProSer 6
|||||
38 GGCACCTCTGCTCCGTCA 21

seq_name: gb_est1:AW847482

seq_documentation_block:
LOCUS      AW847482      121 bp      mRNA      linear      EST 19-MAY-2000
DEFINITION RC6-CT0209-031099-021-c04 CT0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW847482
VERSION    AW847482.1 GI:7942999
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 121)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC6-CT0209-031
            099-021-c04&st3=1999-10-03&st4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 121.
            Location/Qualifiers
                1..121
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="CT0209"
                /dev_stage="Adult"
            /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
            SmaI; A mini-library was made by cloning products derived
            from ORESTES PCR (U.S. Letters Patent application No. 196
            ,716 - Ludwig Institute for Cancer Research) profiles
            into the puc 18 vector. Reverse transcription of tissue
            mRNA and cDNA amplification were performed under low
            stringency conditions."
            BASE COUNT      36 a      41 c      18 g      26 t
            ORIGIN
            alignment_scores:

```

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AW847482 ..

Align seg 1/1 to: AW847482 from: 1 to: 121

1 GlyThr*****ProSer 6
 |||||
 89 GGTACAAGCTCCCTCC 106

seq_name: gb_est1:AA065904

seq_documentation_block:

LOCUS AA065904 122 bp mRNA linear EST 03-FEB-1997
 DEFINITION ml50e11.r1 Stratagene mouse testis (#937308) Mus musculus cDNA
 clone IMAGE:515468 5', mRNA sequence.

ACCESSION AA065904

VERSION AA065904.1 GI:1562634

KEYWORDS

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 122)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HMI Mouse EST Project

JOURNAL

COMMENT

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:309316

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 110.

FEATURES

source

1..122
 /organism="Mus musculus"
 /strain="Inbred CO-1"
 /db_xref="taxon:10090"
 /clone="IMAGE:515468"
 /clone_lib="Stratagene mouse testis (#937308)"
 /sex="males"
 /tissue_type="testis"
 /dev_stage="10-12 week old"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: testis; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
 -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
 sequence: 5' CTCGAGTTTATTTT 3'"

BASE COUNT

ORIGIN 19 a 37 c 37 g 29 t

alignment_scores:

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x AA065904/rev ..

Align seg 1/1 to reverse of: AA065904 from: 1 to: 122

1 GlyThr*****ProSer 6
 |||||
 84 GGCACAGCAGCTCCATCA 67

seq_name: gb_est2:BF856188

seq_documentation_block:

LOCUS BF856188 122 bp mRNA linear EST 16-JAN-2001
 DEFINITION CM4-FN0195-251100-468-c06 FN0195 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF856188

VERSION BF856188.1 GI:12243932

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 122)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

20202663

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-FN0195-251100-468-c06&t3=2000-11-25&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 122.

FEATURES

source

1..122
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="FN0195"
 /dev_stage="Adult"
 /note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI
 ; Site_2: SmaI; A mini-library was made by cloning
 products derived from ORESTES PCR (U.S. Letters Patent
 application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the puc 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."

BASE COUNT

ORIGIN 24 a 43 c 30 g 25 t

alignment_scores:

Quality: 22.00 Length: 6
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

BASK-853-CLAIM5 x BF856188 ..

Align seg 1/1 to: BF856188 from: 1 to: 122

1 GlyThr*****ProSer 6

```

|||||
78 GGAACATCAGCTCCCTCC 95

seq_name: gb_est2:BI777877

seq_documentation_block:
LOCUS      BI777877
DEFINITION Ebro07_SQ001_B03_R IGF Barley Ebro07 library Hordeum vulgare cDNA
clone Ebro07_SQ001_B03 5', mRNA sequence.
ACCESSION  BI777877
VERSION     BI777877.1
KEYWORDS   EST.
SOURCE     Hordeum vulgare
ORGANISM   Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 122)
AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
TITLE     Development of Barley Transcriptome Resources
JOURNAL   Unpublished (2001)
COMMENT   Contact: Waugh R
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse
FEATURES             Location/Qualifiers
     source           1..122
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="Ebro07_SQ001_B03"
                     /clone_lib="IGF Barley Ebro07 library"
                     /tissue_type="Etiolated root and shoot"
                     /lab_host="DH10B"
                     /notes="Vector: pSPORT1; Site.1: Sal I; Site.2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from shoot and root material of 3 week old
etiolated barley plants. Developed as part of the barley
transcriptome resources of BBSRC/SEERAD funded cereal IGF
(investigating Gene Function) project."
BASE COUNT          17 a 50 c 37 g 18 t
ORIGIN
alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x BI777877/rev
Align seg 1/1 to reverse of: BI777877 from: 1 to: 122
1 GlyThr*****ProSer 6
|||||
113 GGAACGACGTCGCTCTCT 96

seq_name: gb_est1:AW878149
seq_documentation_block:
LOCUS      AW878149
DEFINITION MR3-OT0007-100300-102-a05 OT0007 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW878149
VERSION     AW878149.1
KEYWORDS   EST.
SOURCE     human.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 123)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20203663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-MR3-OT0007-100
300-102-a05st3-2000-03-10st4-1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 122.
FEATURES             Location/Qualifiers
     source           1..123
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="OT0007"
                     /dev_stage="Adult"
                     /note="Organ: ovary; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT          21 a 34 c 29 g 39 t
ORIGIN
alignment_scores:
    Quality: 22.00      Length: 6
    Ratio: 3.667       Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x AW878149
Align seg 1/1 to: AW878149 from: 1 to: 123
1 GlyThr*****ProSer 6
|||||
17 GGCACCTCTCTCTCTCT 34

seq_name: gb_est2:N60943
seq_documentation_block:
LOCUS      N60943
DEFINITION TGESTzy14g01.r1 TgrH Tachyzoite cDNA Toxoplasma gondii cDNA clone
tzy14g01.r1 5', mRNA sequence.
ACCESSION  N60943
VERSION     N60943.1
KEYWORDS   EST.
SOURCE     Toxoplasma gondii.
ORGANISM   Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.

```

REFERENCE
1 (bases 1 to 123)
AUTHORS
Hehl, A., Manger, I., Marra, M., Sibley, L.D., Ajioka, J.A., Aslett, M.A.,
Dietrich, N., Dubuque, T., Hillier, L., Kucaba, T., Wan, K.L.,
Waterston, R.H. and Boothroyd, J.
TITLE
WashU-Merck-Stanford-NIH Toxoplasma EST project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@watson.wustl.edu
David Sibley at toxoest@borcim.wustl.edu for further information
relating to organism, clone or library availability.
Seq primer: T3
High quality sequence stop: 97.
FEATURES
Location/Qualifiers
source
1..123
/organism="Toxoplasma gondii"
/strain="RH"
/db_xref="taxon:5811"
/clone_lib="TgrH Tachyzoite cDNA"
/lab_host="XLI-Blue MRF"
/note="Vector: Lambda ZAP; Site_1: EcoRI; Site_2: XhoI;
Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
constructed by K.L. Wan, Cambridge University. cDNAs were
synthesized from polyA RNAs by oligo d(T) priming and
directionally cloned into the EcoRI to XhoI sites of the
Lambda ZapII vector using the ZAP-cDNA synthesis kit
(Stratagene). WARNING: the library contains a small
percentage of cDNAs derived from the human host cells."
BASE COUNT
49 a 23 c 30 g 17 t 4 others
ORIGIN
alignment_scores:
Quality: 22.00 Length: 6
Ratio: 3.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
BASK-853-CLAIM5 x N60943 ..
Align seg 1/1 to: N60943 from: 1 to: 123
1 GlyThr*****ProSer 6
|||||
1 GGCACGACGAGCTCCACG 18